

Chapter 14

Metric spaces and continuity 1

Introduction

In Chapter 13, *Distance and continuity*, we saw how the sequential definition of continuity for functions from $\mathbb{R}$ to $\mathbb{R}$ can be used as a basis for a definition of continuity for functions from $\mathbb{R}^2$ to $\mathbb{R}$. In this chapter we continue our journey towards understanding continuity in a general setting by finding a definition of continuity for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$, which is inspired by our earlier definition of continuity for functions from $\mathbb{R}^2$ to $\mathbb{R}$. The *Euclidean distance* between pairs of points in $\mathbb{R}^n$ (and $\mathbb{R}^m$) plays a key role in this definition. In Chapter 13 we found that the Euclidean distance function for the plane has three properties: non-negativity, symmetry and the Triangle Inequality. In a similar way, we find in this chapter that the Euclidean distance on $\mathbb{R}^n$ satisfies the same three properties. Our creative leap is to use this observation to raise these properties to the status of axioms and to say that any function that satisfies them is a well-defined distance function, known as a *metric*.

In Section 2, we start to explore the implications of this definition and see how the notion of a metric allows our definitions of convergence and continuity in Euclidean spaces to be extended naturally to other spaces. We also see how the concept of a metric leads to different ways of measuring ‘distance’ on the same underlying set of points.

Chapter guide

Section 1 of this chapter to some extent repeats the work of the previous chapter, but in the context of $\mathbb{R}^n$ rather than the plane. Apart from some extra notational complexity, you will already be familiar with the ideas.

The key idea in the chapter is the decision to use three particular properties of the Euclidean distance as the basis for defining what we mean by a general distance function, a *metric*. This process of using the properties of the (reasonably) familiar as the basis of a definition for a whole class of objects can be challenging when first encountered, and it is likely that you will need to spend a significant proportion of your time studying Section 2.

Sections 3 and 4 are very brief. Section 3 develops the idea of sequences and convergence in metric spaces, before Section 4 extends the definition of continuity from Euclidean spaces to metric spaces.

1 Continuity of functions from $\mathbb{R}^n$ to $\mathbb{R}^m$

In this section we continue our process of abstracting the ideas of distance and the continuity of functions. Having dealt with distance and continuity in one and two dimensions, we now look at them in higher dimensions; in particular, we examine distance in $\mathbb{R}^n$ (for general n) and the continuity of functions from $\mathbb{R}^n$ to $\mathbb{R}^m$.

If we look back at our work in the last chapter, we see that there are several notions that played an important role in reaching a definition of continuity for functions from $\mathbb{R}^2 \rightarrow \mathbb{R}$: we needed to be able to measure distance between pairs of points in both $\mathbb{R}$ and $\mathbb{R}^2$, and we also needed a notion of convergence for sequences in both $\mathbb{R}$ and $\mathbb{R}^2$. The methods we adopt for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ are inspired by those we used for functions from $\mathbb{R}^2$ to $\mathbb{R}$ but with some extra notational complication because of the need to track the dimensions of the spaces involved. Since we cannot usually easily visualise distances and functions in such spaces, we rely on algebraic methods to carry through these extensions.

However, before we begin this process, we digress briefly to explain what we mean by A^k when A is a set and $k \in \mathbb{N}$.

Replacing $\mathbb{R}^n$ by $\mathbb{R}^3$ can sometimes help you to understand general results.

The product set is sometimes called the *Cartesian product*.

Definition 1.1 Product set

Let $A_1, A_2, \dots, A_k$ be sets. The **product set** $A_1 \times A_2 \times \dots \times A_k$ is defined to be the set consisting of all ordered k -tuples $(a_1, a_2, \dots, a_k)$ where $a_1 \in A_1, a_2 \in A_2, \dots, a_k \in A_k$. Thus

$$A_1 \times A_2 \times \dots \times A_k = \{(a_1, a_2, \dots, a_k) : a_i \in A_i \text{ for } i = 1, \dots, k\}.$$

If $A_1 = A_2 = \dots = A_k = A$, then the product set $A_1 \times A_2 \times \dots \times A_k$ is written simply as A^k and

$$A^k = \{(a_1, a_2, \dots, a_k) : a_i \in A \text{ for } i = 1, \dots, k\}.$$

Remarks

1. The term *product* is used for this set because, when $A_1, A_2, \dots, A_k$ are finite sets with $m_1, m_2, \dots, m_k$ elements respectively, the set $A_1 \times A_2 \times \dots \times A_k$ has $m_1 m_2 \dots m_k$ elements. For example, if $A_1 = \{1, 2\}$ and $A_2 = \{1, 4, 9\}$, then

$$A_1 \times A_2 = \{(1, 1), (1, 4), (1, 9), (2, 1), (2, 4), (2, 9)\},$$

which has $2 \times 3 = 6$ elements.

2. If $k = 2$ and $A_1 = A_2 = \mathbb{R}$, then $A_1 \times A_2 = \mathbb{R} \times \mathbb{R}$, the set of all ordered pairs of real numbers. Thus $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$. Similarly, if $A_1, A_2, \dots, A_k$ are all equal to $\mathbb{R}$, then the product set $A_1 \times A_2 \times \dots \times A_k = \mathbb{R}^k$.

3. The definition can be extended to the product of infinitely many sets. In particular, the product of infinitely many copies of $\mathbb{R}$ gives the set $\mathbb{R}^\infty$, the set of all ordered infinite tuples of real numbers of the form $(a_1, a_2, a_3, \dots)$. This is equivalent to the set of all real sequences. We can write these infinite tuples in the form $(a_n)_{n=1}^\infty$, which coincides with our notation for sequences.
4. We often denote an element of a product set $A_1 \times A_2 \times \dots \times A_k$ as $\mathbf{a} = (a_1, a_2, \dots, a_k)$ and we say that two elements $\mathbf{a} = (a_1, a_2, \dots, a_k)$ and $\mathbf{b} = (b_1, b_2, \dots, b_k)$ are equal if $a_i = b_i$ for $i = 1, 2, \dots, k$.

This is one reason why we use (a_n) as our notation for sequences, since parentheses imply an ordering whereas braces do not.

Exercise 1.1

Write down all the elements of $A_1 \times A_2$ when $A_1 = \{a, b\}$ and $A_2 = \{0, 1, 2, 3\}$.

1.1 Functions and distance

In order to fix notation, we start by discussing briefly points, functions and our usual ideas of distance in $\mathbb{R}^n$, n -dimensional Euclidean space.

Points in $\mathbb{R}^n$ have n coordinates, and are denoted by ordered n -tuples $(x_1, x_2, \dots, x_n)$. As we did with $\mathbb{R}^2$, we use the notation $\mathbf{x}$ for the point $(x_1, x_2, \dots, x_n)$ and $\mathbf{0}$ for the point $(0, 0, \dots, 0)$, the **origin** of $\mathbb{R}^n$. Also, if $\mathbf{a} = (a_1, a_2, \dots, a_n)$ and $\mathbf{b} = (b_1, b_2, \dots, b_n)$, then the sum and difference of $\mathbf{a}$ and $\mathbf{b}$ are given by

$$\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, \dots, a_n + b_n)$$

and

$$\mathbf{a} - \mathbf{b} = (a_1 - b_1, a_2 - b_2, \dots, a_n - b_n)$$

respectively.

Functions between Euclidean spaces will have domain and codomain both Euclidean but possibly of different dimension. Some examples of such functions are:

- $f: \mathbb{R}^4 \rightarrow \mathbb{R}$ given by $f(x_1, x_2, x_3, x_4) = x_1 + x_2 + x_3 + x_4$, so that, for instance, $f(1, -4, 0, 2) = -1$
- $g: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ given by $g(x_1, x_2, x_3) = (x_3, x_2)$, so that, for instance, $g(3, 7, 1) = (1, 7)$
- $h: \mathbb{R}^5 \rightarrow \mathbb{R}^3$ given by $h(x_1, x_2, x_3, x_4, x_5) = (2x_1, x_2^2 + 3x_4, x_3 + x_5)$, so that, for instance, $h(1, -2, 3, -3, 4) = (2, -5, 7)$
- $k: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ given by $k(x_1, x_2) = (\cos 2\pi x_1, \sin \pi x_2, x_1 \sin 2\pi x_2)$, so that, for instance, $k(1, 1.5) = (1, -1, 0)$.

Notice that the domains and codomains of these functions can have different dimensions. In general, we consider functions from $\mathbb{R}^n$ to $\mathbb{R}^m$.

In the previous chapter we used the definition of continuity for functions from $\mathbb{R}$ to $\mathbb{R}$ as a basis to define what it meant for functions from $\mathbb{R}^2$ to $\mathbb{R}$ to be continuous. A fundamental observation in this process was that the modulus function acts as the distance function on $\mathbb{R}$; this then led us to consider the distance function $d^{(2)}$ on the plane. In order to define continuity for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$, we need to be able to measure the distance between pairs of points in $\mathbb{R}^n$ (and pairs of points in $\mathbb{R}^m$).

Definition 1.2 Euclidean distance on $\mathbb{R}^n$

The **Euclidean distance** between points $\mathbf{a} = (a_1, a_2, \dots, a_n)$ and $\mathbf{b} = (b_1, b_2, \dots, b_n)$ in $\mathbb{R}^n$ is given by the formula

$$d^{(n)}(\mathbf{a}, \mathbf{b}) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \dots + (b_n - a_n)^2}.$$

When it is clear from the context which dimension of Euclidean space we are working with, we may simply write $d(\mathbf{a}, \mathbf{b})$ for this distance.

Remarks

1. When $n = 1$ and $n = 2$, we recover the Euclidean distances in $\mathbb{R}$ and $\mathbb{R}^2$, respectively: when $n = 1$, we obtain $d^{(1)}(a, b) = \sqrt{(b - a)^2} = |b - a|$, and when $n = 2$, we obtain the distance between points in the plane.
2. As with $d^{(2)}$, we have $d^{(n)}(\mathbf{a}, \mathbf{b}) = d^{(n)}(\mathbf{a} - \mathbf{b}, \mathbf{0}) = d^{(n)}(\mathbf{b} - \mathbf{a}, \mathbf{0})$.

For example, in $\mathbb{R}^4$,

$$\begin{aligned} d^{(4)}((1, 2, 3, 4), (2, -1, 0, 3)) &= \sqrt{(2 - 1)^2 + (-1 - 2)^2 + (0 - 3)^2 + (3 - 4)^2} \\ &= \sqrt{1 + 9 + 9 + 1} = \sqrt{20} = 2\sqrt{5}. \end{aligned}$$

Exercise 1.2

In $\mathbb{R}^5$, calculate the distance $d^{(5)}((4, 2, 3, 0, -1), (2, -3, -1, 0, 2))$.

1.2 Properties of the Euclidean distance in $\mathbb{R}^n$

In Subsection 3.4 of Chapter 13 we gave three properties possessed by the Euclidean distance functions on $\mathbb{R}$ and on $\mathbb{R}^2$. In this subsection we verify that the Euclidean distance function $d^{(n)}(\mathbf{a}, \mathbf{b})$ on $\mathbb{R}^n$ has the same properties.

The first property is that distances are never negative and that the distance between two points is zero if, and only if, the two points are the same.

(M1) For each $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, we have $d^{(n)}(\mathbf{a}, \mathbf{b}) \geq 0$, with $d^{(n)}(\mathbf{a}, \mathbf{b}) = 0$ if, and only if, $\mathbf{a} = \mathbf{b}$.

This is referred to as the *non-negativity property*.

The property that distances are never negative is built into $d^{(n)}$ through the square root. To show that the second part of (M1) holds, suppose first that $\mathbf{a} = \mathbf{b}$. Then $a_i = b_i$ for all values of i , and so each term $(b_i - a_i)^2$ inside the square root of the Euclidean distance function

$$d^{(n)}(\mathbf{a}, \mathbf{b}) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \cdots + (b_n - a_n)^2}$$

is 0; hence $d^{(n)}(\mathbf{a}, \mathbf{b}) = 0$. Conversely, since each term $(b_i - a_i)^2$ in $d^{(n)}(\mathbf{a}, \mathbf{b})$ is non-negative, when $d^{(n)}(\mathbf{a}, \mathbf{b}) = 0$ each term inside the square root must be equal to 0, and so $a_i = b_i$ for all values of i . Thus $\mathbf{a} = \mathbf{b}$. Therefore, the distance $d^{(n)}(\mathbf{a}, \mathbf{b})$ between two points is 0 *if, and only if*, $\mathbf{a} = \mathbf{b}$.

The next property is that the distance from $\mathbf{a}$ to $\mathbf{b}$ is the same as the distance from $\mathbf{b}$ to $\mathbf{a}$.

(M2) For each $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, we have $d^{(n)}(\mathbf{a}, \mathbf{b}) = d^{(n)}(\mathbf{b}, \mathbf{a})$.

This is referred to as the *symmetry property*.

Exercise 1.3

Verify property (M2) by examining $d^{(n)}(\mathbf{a}, \mathbf{b})$ and $d^{(n)}(\mathbf{b}, \mathbf{a})$ for $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$.

The symmetry property of $d^{(n)}(\mathbf{a}, \mathbf{b})$ allows us to use the following two forms interchangeably:

$$d^{(n)}(\mathbf{a}, \mathbf{b}) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \cdots + (b_n - a_n)^2},$$

$$d^{(n)}(\mathbf{a}, \mathbf{b}) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}.$$

The third property is the **Triangle Inequality**.

(M3) For each $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$, we have $d^{(n)}(\mathbf{a}, \mathbf{c}) \leq d^{(n)}(\mathbf{a}, \mathbf{b}) + d^{(n)}(\mathbf{b}, \mathbf{c})$.

There are many different proofs of the Triangle Inequality. We defer a proof of it to Subsection 1.5, where we will give one that is purely algebraic.

We summarise these results in the following proposition.

Proposition 1.3 *Properties of Euclidean distance on $\mathbb{R}^n$*

The Euclidean distance function $d^{(n)}$ on $\mathbb{R}^n$ has the following properties.

For each $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$:

$$(M1) \quad d^{(n)}(\mathbf{a}, \mathbf{b}) \geq 0, \text{ with equality holding if, and only if, } \mathbf{a} = \mathbf{b}$$

$$(M2) \quad d^{(n)}(\mathbf{a}, \mathbf{b}) = d^{(n)}(\mathbf{b}, \mathbf{a})$$

$$(M3) \quad d^{(n)}(\mathbf{a}, \mathbf{c}) \leq d^{(n)}(\mathbf{a}, \mathbf{b}) + d^{(n)}(\mathbf{b}, \mathbf{c}).$$

non-negativity

symmetry

Triangle Inequality

In Section 2, we will use these three properties as the definition for a general distance function, so you need to be familiar with them. In particular, be sure that you understand the meaning of the Triangle Inequality, as it is of basic importance.

The Triangle Inequality for $d^{(n)}$, like that for distances in $\mathbb{R}$, has a ‘reversed’ form.

Proposition 1.4 *Reverse Triangle Inequality for $\mathbb{R}^n$*

The Euclidean distance function $d^{(n)}$ on $\mathbb{R}^n$ has the following property.

$$(M3a) \quad \text{If } \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n, \text{ then } d^{(n)}(\mathbf{b}, \mathbf{c}) \geq |d^{(n)}(\mathbf{a}, \mathbf{c}) - d^{(n)}(\mathbf{a}, \mathbf{b})|.$$

Proof From (M3) we obtain, for each $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$,

$$d^{(n)}(\mathbf{a}, \mathbf{c}) \leq d^{(n)}(\mathbf{a}, \mathbf{b}) + d^{(n)}(\mathbf{b}, \mathbf{c}).$$

Subtracting $d^{(n)}(\mathbf{a}, \mathbf{b})$ from each side, we obtain

$$d^{(n)}(\mathbf{a}, \mathbf{c}) - d^{(n)}(\mathbf{a}, \mathbf{b}) \leq d^{(n)}(\mathbf{b}, \mathbf{c}). \quad (1)$$

Since $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are arbitrary points in $\mathbb{R}^n$, we can freely interchange their roles in the Triangle Inequality. Thus, interchanging the roles of $\mathbf{b}$ and $\mathbf{c}$ in equation (1), we get

$$d^{(n)}(\mathbf{a}, \mathbf{b}) - d^{(n)}(\mathbf{a}, \mathbf{c}) \leq d^{(n)}(\mathbf{c}, \mathbf{b}) = d^{(n)}(\mathbf{b}, \mathbf{c}), \text{ by (M2).}$$

This can be rearranged to give

$$-d^{(n)}(\mathbf{b}, \mathbf{c}) \leq d^{(n)}(\mathbf{a}, \mathbf{c}) - d^{(n)}(\mathbf{a}, \mathbf{b}). \quad (2)$$

Combining inequalities (1) and (2), we obtain

$$-d^{(n)}(\mathbf{b}, \mathbf{c}) \leq d^{(n)}(\mathbf{a}, \mathbf{c}) - d^{(n)}(\mathbf{a}, \mathbf{b}) \leq d^{(n)}(\mathbf{b}, \mathbf{c}),$$

which is equivalent to

$$|d^{(n)}(\mathbf{a}, \mathbf{c}) - d^{(n)}(\mathbf{a}, \mathbf{b})| \leq d^{(n)}(\mathbf{b}, \mathbf{c}),$$

as required. ■

We finish this subsection by examining the domain and codomain of the distance function determined by the Euclidean distance for $\mathbb{R}^n$. The Euclidean distance formula

$$d^{(n)}(\mathbf{a}, \mathbf{b}) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \cdots + (b_n - a_n)^2}$$

defines a function of two variables $\mathbf{a}$ and $\mathbf{b}$, each from the set $\mathbb{R}^n$, which produce a (non-negative) real number, the distance.

The domain of the distance function $d^{(1)}$ is simply the set of all ordered pairs (a, b) where $a \in \mathbb{R}$ and $b \in \mathbb{R}$ – that is, the domain is the product set $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$. However, each of the arguments $\mathbf{a} = (a_1, a_2)$ and $\mathbf{b} = (b_1, b_2)$ of the distance function $d^{(2)}$ is itself an ordered pair of real numbers from $\mathbb{R}^2$. Thus the domain of the distance function $d^{(2)}$ is the set of all ordered pairs of elements of $\mathbb{R}^2$, i.e. the product set $\mathbb{R}^2 \times \mathbb{R}^2$. In the general case, the domain of the distance function $d^{(n)}$ is the product set $\mathbb{R}^n \times \mathbb{R}^n$.

For each n , we take the codomain of $d^{(n)}$ to be $\mathbb{R}$.

Hence $d^{(n)}$ is a function

$$d^{(n)}: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

from the set of all ordered pairs of ordered n -tuples of real numbers to the set of real numbers.

We could take the codomain to be any set that contains $[0, \infty)$.

1.3 Sequences in $\mathbb{R}^n$

In the previous chapter we studied real sequences and sequences of points in $\mathbb{R}^2$. We now extend these ideas to include sequences in $\mathbb{R}^n$. Following Chapter 13, we will use convergence of sequences (in $\mathbb{R}^n$ and $\mathbb{R}^m$) to define the continuity of functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ (in the next subsection).

The definitions of sequences and convergence in $\mathbb{R}^n$ are very similar to those in the plane: we simply change the set from $\mathbb{R}^2$ to $\mathbb{R}^n$ and replace the Euclidean distance $d^{(2)}$ with the n -dimensional version $d^{(n)}$.

Definition 1.5 Sequence in $\mathbb{R}^n$

A **sequence** in $\mathbb{R}^n$ is an unending ordered list of points in $\mathbb{R}^n$:

$$\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots$$

The point $\mathbf{a}_k = (a_{1,k}, a_{2,k}, \dots, a_{n,k})$ is the **k th term** of the sequence, and the whole sequence is denoted by $(\mathbf{a}_k)$, $(\mathbf{a}_k)_{k=1}^\infty$ or $(\mathbf{a}_k)_{k \in \mathbb{N}}$.

Here we use the letter k for a general term in the sequence as the letter n has already been used to denote the dimension of $\mathbb{R}^n$.

Definition 1.6 Convergence in $\mathbb{R}^n$

A sequence $(\mathbf{a}_k)$ in $\mathbb{R}^n$ **converges** to $\ell \in \mathbb{R}^n$ if $(d^{(n)}(\mathbf{a}_k, \ell))$ is a (real) null sequence.

We say that ℓ is the **limit** of the sequence $(\mathbf{a}_k)$ and write $\mathbf{a}_k \rightarrow \ell$ as $k \rightarrow \infty$, $\lim_{k \rightarrow \infty} \mathbf{a}_k = \ell$ or simply $\mathbf{a}_k \rightarrow \ell$.

A sequence in $\mathbb{R}^n$ that does not converge to any point in $\mathbb{R}^n$ is said to be **divergent**.

Remarks

1. A sequence in $\mathbb{R}^n$ that converges to $\mathbf{0}$ is called a **null sequence**.
2. Note that $(d^{(n)}(\mathbf{a}_k, \ell))_{k \in \mathbb{N}}$ is a sequence of real numbers. Hence, testing whether a given sequence $(\mathbf{a}_k)$ in $\mathbb{R}^n$ converges to $\ell \in \mathbb{R}^n$ reduces to testing whether $(d^{(n)}(\mathbf{a}_k, \ell))$ is a *real* null sequence.
3. Definition 1.6 is equivalent to the following:

A sequence $(\mathbf{a}_k)$ in $\mathbb{R}^n$ converges to ℓ if, for each $\varepsilon > 0$, there is a $K \in \mathbb{N}$ such that

$$d^{(n)}(\mathbf{a}_k, \ell) < \varepsilon \quad \text{whenever } k > K.$$

There is a very important fact about convergent sequences in $\mathbb{R}^n$, which we will prove now before we consider some examples: limits of convergent sequences are unique.

Lemma 1.7 Uniqueness of limits in $\mathbb{R}^n$

Let $(\mathbf{a}_k)$ be a convergent sequence in $\mathbb{R}^n$. If $\mathbf{a}_k \rightarrow \mathbf{r}$ and $\mathbf{a}_k \rightarrow \mathbf{s}$ as $k \rightarrow \infty$, then $\mathbf{r} = \mathbf{s}$.

Proof We let d denote the Euclidean distance on $\mathbb{R}^n$ and show that $d(\mathbf{r}, \mathbf{s}) = 0$. Property (M1) for d then implies that $\mathbf{r} = \mathbf{s}$, which proves the lemma.

The definition of convergence for sequences in $\mathbb{R}^n$ tells us that $(d(\mathbf{a}_k, \mathbf{r}))$ and $(d(\mathbf{a}_k, \mathbf{s}))$ are both *real* null sequences. Hence we can use property (M3) for d , the Triangle Inequality, to estimate that for each $k \in \mathbb{N}$,

$$\begin{aligned} d(\mathbf{r}, \mathbf{s}) &\leq d(\mathbf{r}, \mathbf{a}_k) + d(\mathbf{a}_k, \mathbf{s}) \\ &= d(\mathbf{a}_k, \mathbf{r}) + d(\mathbf{a}_k, \mathbf{s}), \text{ by property (M2) for } d. \end{aligned}$$

However, both $d(\mathbf{a}_k, \mathbf{r}) \rightarrow 0$ and $d(\mathbf{a}_k, \mathbf{s}) \rightarrow 0$ as $k \rightarrow \infty$ and so

$$d(\mathbf{r}, \mathbf{s}) \leq 0.$$

But it is certainly the case (by property (M1) for d) that $d(\mathbf{r}, \mathbf{s}) \geq 0$ and so $d(\mathbf{r}, \mathbf{s}) = 0$, as required. ■

Notice that the proof of this lemma only uses the fact that the Euclidean distance on $\mathbb{R}^n$ has properties (M1)–(M3).

Worked Exercise 1.8

Prove that the sequence $(\mathbf{b}_k)$ in $\mathbb{R}^3$ given by $\mathbf{b}_k = (\frac{k+1}{k}, \frac{(-1)^k}{k}, 3)$ for each $k \in \mathbb{N}$ converges, and find its limit.

Solution

Since $\frac{k+1}{k} \rightarrow 1$ and $\frac{(-1)^k}{k} \rightarrow 0$ as $k \rightarrow \infty$, the terms of $(\mathbf{b}_k)$ appear to get near to $(1, 0, 3)$ as k gets large. So letting d denote the Euclidean distance on $\mathbb{R}^3$, then for $k \in \mathbb{N}$, we calculate

$$\begin{aligned} d(\mathbf{b}_k, (1, 0, 3)) &= \sqrt{\left(1 - \frac{k+1}{k}\right)^2 + \left(0 - \frac{(-1)^k}{k}\right)^2 + (3 - 3)^2} \\ &= \sqrt{\frac{1}{k^2} + \frac{1}{k^2}} = \frac{\sqrt{2}}{k}. \end{aligned}$$

Since $\left(\frac{\sqrt{2}}{k}\right)$ is a multiple of a basic null sequence, we conclude that $(d(\mathbf{b}_k, (1, 0, 3)))$ is a real null sequence and so $(\mathbf{b}_k)$ converges to $(1, 0, 3)$ as $k \rightarrow \infty$ in $\mathbb{R}^3$.

Exercise 1.4

(a) Let the sequence $(\mathbf{a}_k)$ in $\mathbb{R}^4$ be defined by

$$\mathbf{a}_k = \left(\frac{-1}{(-3)^k}, \frac{3^k + 5 \cos k\pi}{3^k}, \sin k\pi, -1 - \frac{2}{3^k} \right).$$

Show that $(\mathbf{a}_k)$ converges, and find its limit.

(b) Prove that any constant sequence in $\mathbb{R}^n$ converges.

If $(\mathbf{a}_k)$ is a sequence in $\mathbb{R}^n$, then the n sequences $(a_{j,k})$ for $j = 1, 2, \dots, n$ can be thought of as individual sequences of real numbers in their own right. We refer to them as **coordinate sequences**. The next proposition makes clear the connection between the convergence of a sequence $(\mathbf{a}_k)$ in $\mathbb{R}^n$ and the convergence of these n individual coordinate sequences.

Proposition 1.9 extends the result of Theorem 3.7 from Chapter 13.

Proposition 1.9

Let $(\mathbf{a}_k)$ be a sequence in $\mathbb{R}^n$, with $\mathbf{a}_k = (a_{1,k}, \dots, a_{n,k})$, and let $\ell = (\ell_1, \dots, \ell_n) \in \mathbb{R}^n$. Then

$$\mathbf{a}_k \rightarrow \ell \quad \text{as } k \rightarrow \infty$$

if, and only if,

$$a_{j,k} \rightarrow \ell_j \quad \text{as } k \rightarrow \infty, \text{ for each } j = 1, 2, \dots, n.$$

Proposition 1.9 is useful because it tells us that to test the convergence of a given sequence $(\mathbf{a}_k)$ in $\mathbb{R}^n$, it is enough to test the convergence of the real coordinate sequences $(a_{j,k})$ for each j . We give some applications of this proposition before we ask you to prove it in Exercise 1.7.

Worked Exercise 1.10

Use Proposition 1.9 to prove the Sum Rule for convergent sequences in $\mathbb{R}^n$: if $\mathbf{a}_k \rightarrow \ell$ and $\mathbf{b}_k \rightarrow \mathbf{m}$ as $k \rightarrow \infty$, then $\mathbf{a}_k + \mathbf{b}_k \rightarrow \ell + \mathbf{m}$ as $k \rightarrow \infty$.

Solution

Let $(\mathbf{a}_k)$ and $(\mathbf{b}_k)$ be sequences in $\mathbb{R}^n$ such that $\mathbf{a}_k \rightarrow \ell$ and $\mathbf{b}_k \rightarrow \mathbf{m}$. We need to show that $(\mathbf{a}_k + \mathbf{b}_k)$ converges to $\ell + \mathbf{m}$.

Proposition 1.9 gives that

$$a_{j,k} \rightarrow \ell_j \quad \text{and} \quad b_{j,k} \rightarrow m_j \quad \text{as } k \rightarrow \infty, \text{ for } j = 1, 2, \dots, n.$$

By applying the Sum Rule for convergent real sequences, we see that

$$a_{j,k} + b_{j,k} \rightarrow \ell_j + m_j \quad \text{as } k \rightarrow \infty, \text{ for } j = 1, 2, \dots, n.$$

Using Proposition 1.9 again now shows that $\mathbf{a}_k + \mathbf{b}_k \rightarrow \ell + \mathbf{m}$ as $k \rightarrow \infty$.

Exercise 1.5

Prove the Multiple Rule for a convergent sequence $\mathbf{a}_k \rightarrow \ell$ in $\mathbb{R}^n$. That is, show that $\lambda \mathbf{a}_k \rightarrow \lambda \ell$ for any $\lambda \in \mathbb{R}$.

For $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$, we write $\lambda \mathbf{x} = (\lambda x_1, \dots, \lambda x_n)$.

Exercise 1.6

Use Proposition 1.9 to show that the sequence $(\mathbf{a}_k)$ in $\mathbb{R}^3$ given by

$$\mathbf{a}_k = \left(\frac{k^3 + 3^k}{k!}, 1 - \frac{1}{\sqrt{k}}, 2^{-k}(5 - 2^k) \right)$$

converges, and find its limit.

Exercise 1.7

Adapt the proof of Theorem 3.7 from Chapter 13 to prove Proposition 1.9.

For convenience, we summarise the combination rules that we have derived in these exercises.

Combination rules for convergent sequences in $\mathbb{R}^n$

If $\mathbf{a}_k \rightarrow \ell$ and $\mathbf{b}_k \rightarrow \mathbf{m}$ as $k \rightarrow \infty$, then:

Sum Rule $\mathbf{a}_k + \mathbf{b}_k \rightarrow \ell + \mathbf{m}$ as $k \rightarrow \infty$

Multiple Rule $\lambda \mathbf{a}_k \rightarrow \lambda \ell$ as $k \rightarrow \infty$, for any $\lambda \in \mathbb{R}$.

We do not give a Product Rule for sequences in $\mathbb{R}^n$ because we have not defined what it means to ‘multiply’ points of $\mathbb{R}^n$ together.

Proposition 1.9 also provides a very convenient method for showing that a sequence in $\mathbb{R}^n$ diverges. It tells us that a sequence $(\mathbf{a}_k) = ((a_{1,k}, \dots, a_{n,k}))$ is divergent if *any one* of the coordinate sequences $(a_{j,k})$ is divergent. For example, the sequence in $\mathbb{R}^3$ given by $\mathbf{a}_k = (\frac{1}{k}, \frac{2k^2+1}{k^2}, (-1)^k)$ diverges because the third coordinate sequence $(a_{3,k}) = ((-1)^k)$ does not converge to any real number.

The sequence $((-1)^k)$ was shown to be divergent in Worked Exercise 1.26 of Chapter 13.

Exercise 1.8

Let $(\mathbf{b}_k)$ be the sequence in $\mathbb{R}^4$ given by $\mathbf{b}_k = \left(\frac{k!}{k^k}, k, \frac{2}{\sqrt{k}}, k^{1/\log k}\right)$.

Determine whether $(\mathbf{b}_k)$ converges, justifying your answer.

1.4 Continuity for functions between Euclidean spaces

We are now ready to extend the definition of continuity to functions from $\mathbb{R}^n$ to $\mathbb{R}^m$. Given all our work up to this point, it is not too difficult to see what we have to do to define continuity for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$: we just have to work with sequences in $\mathbb{R}^n$ in the domain and sequences in $\mathbb{R}^m$ in the codomain.

The resulting definition of continuity for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ mimics that for functions from $\mathbb{R}^2$ to $\mathbb{R}$, which was given in Chapter 13.

See Chapter 13, Definition 3.8.

Definition 1.11 *Continuity from $\mathbb{R}^n$ to $\mathbb{R}^m$*

Let $A \subseteq \mathbb{R}^n$ and let $f: A \rightarrow \mathbb{R}^m$ be a function.

Then f is **continuous** at $\mathbf{a} \in A$ if:

whenever $(\mathbf{x}_k)$ is a sequence in A for which $\mathbf{x}_k \rightarrow \mathbf{a}$ as $k \rightarrow \infty$,
then the sequence $(f(\mathbf{x}_k))$ converges to $f(\mathbf{a})$.

If f does not satisfy this condition at $\mathbf{a} \in A$ – that is, there is a sequence $(\mathbf{x}_k)$ in A for which $\mathbf{x}_k \rightarrow \mathbf{a}$ as $k \rightarrow \infty$ but $f(\mathbf{x}_k)$ does not converge to $f(\mathbf{a})$ – then we say that f is **discontinuous** at $\mathbf{a}$.

Note that the sequence $(f(\mathbf{x}_k))$ in the above definition is a sequence in $\mathbb{R}^m$.

Definition 1.12 *Continuity on a set in $\mathbb{R}^n$*

Let $S \subseteq A \subseteq \mathbb{R}^n$ and let $f: A \rightarrow \mathbb{R}^m$ be a function.

We say that f is **continuous** on S when f is continuous at each point in S .

We say that f is **discontinuous** on S if f is discontinuous at at least one point in S .

It is worth reflecting briefly on how we have arrived at this definition of continuity. We first examined the notion of distance, extending the distance function $d^{(1)}(a, b) = |b - a|$ on $\mathbb{R}$ to obtain the Euclidean distance function $d^{(n)}(\mathbf{a}, \mathbf{b})$ on $\mathbb{R}^n$. Next, we used this idea of distance to provide a definition of the convergence of sequences in higher dimensions. Finally, we have used convergent sequences to say what it means for a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ to be continuous. So we see that it is our ideas of distance that underpin the concept of continuity.

We now consider some examples of continuous functions from $\mathbb{R}^n$ to $\mathbb{R}^m$, for various values of n and m .

Worked Exercise 1.13

Prove that the function $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ given by

$$f(x_1, x_2, x_3) = (x_1 + x_3, x_2)$$

is continuous on $\mathbb{R}^3$.

Solution

Let $\mathbf{a}$ be a (general) point of $\mathbb{R}^3$ and let $(\mathbf{x}_k)$ be a sequence in $\mathbb{R}^3$ such that $\mathbf{x}_k \rightarrow \mathbf{a}$ as $k \rightarrow \infty$.

Writing $\mathbf{x}_k = (x_{1,k}, x_{2,k}, x_{3,k})$ and $\mathbf{a} = (a_1, a_2, a_3)$, Proposition 1.9 tells us that

$$x_{1,k} \rightarrow a_1, \quad x_{2,k} \rightarrow a_2 \quad \text{and} \quad x_{3,k} \rightarrow a_3 \quad \text{as } k \rightarrow \infty. \quad (3)$$

We need to show that the sequence $(f(\mathbf{x}_k))$ converges to $f(\mathbf{a})$ in $\mathbb{R}^2$, where

$$f(\mathbf{x}_k) = f(x_{1,k}, x_{2,k}, x_{3,k}) = (x_{1,k} + x_{3,k}, x_{2,k}).$$

Applying the Sum Rule (for real sequences) to equation (3) gives that $x_{1,k} + x_{3,k} \rightarrow a_1 + a_3$ as $k \rightarrow \infty$. Therefore, using equation (3) and Proposition 1.9 again, it follows that

$$f(\mathbf{x}_k) = (x_{1,k} + x_{3,k}, x_{2,k}) \rightarrow (a_1 + a_3, a_2) = f(\mathbf{a}) \quad \text{as } k \rightarrow \infty.$$

We conclude that f is continuous at $\mathbf{a}$.

Since $\mathbf{a}$ is an arbitrary point of $\mathbb{R}^3$, f is continuous on $\mathbb{R}^3$.

Worked Exercise 1.14

Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be given by

$$f(\mathbf{x}) = (\text{int}(2x_1), 3x_2),$$

where, as usual, $\text{int}(a)$ denotes the integer part of a (the largest integer less than or equal to a). Prove that f is a discontinuous function.

Solution

We need to find a point at which the function f is discontinuous. That is, we must find $\mathbf{a} \in \mathbb{R}^2$ and a sequence $(\mathbf{x}_k)$ of points in the plane for which $\mathbf{x}_k \rightarrow \mathbf{a}$ and yet $f(\mathbf{x}_k) \not\rightarrow f(\mathbf{a})$ as $k \rightarrow \infty$.

It turns out that there are many possible choices for $\mathbf{a}$. Consider the point $\mathbf{a} = (\frac{1}{2}, 0)$. Then $f(\mathbf{a}) = (\text{int}(1), 0) = (1, 0)$ but if $\mathbf{x}$ is a point with x_1 -coordinate just a little less than $\frac{1}{2}$, then $\text{int}(2x_1) = 0$. Hence if we take a sequence of points in the plane $(\mathbf{x}_k)$ that converge to $\mathbf{a}$ from the ‘left’ (Figure 1.1), then $f(\mathbf{x}_k)$ will have first coordinate zero for large enough k and so cannot converge to $f(\mathbf{a})$. For example, take $\mathbf{x}_k = (\frac{1}{2} - \frac{1}{2k}, 0)$. Then certainly $\mathbf{x}_k \rightarrow (\frac{1}{2}, 0) = \mathbf{a}$ as $k \rightarrow \infty$ but for each $k \in \mathbb{N}$, $f(\mathbf{x}_k) = (\text{int}(1 - \frac{1}{k}), 0) = (0, 0) \not\rightarrow (1, 0) = f(\mathbf{a})$. Hence f is discontinuous at $\mathbf{a}$.

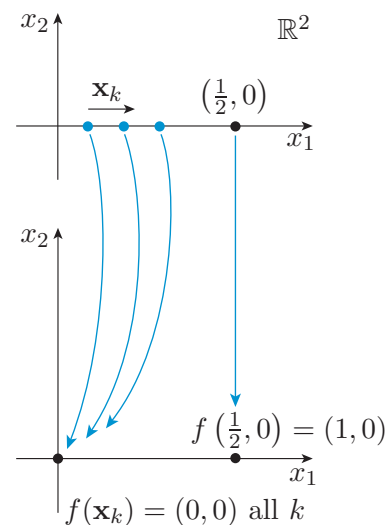


Figure 1.1 Sequence of points converging to $\mathbf{a}$ from the left

In fact f is discontinuous at any point of the form $(\frac{m}{2}, x_2)$ for $m \in \mathbb{Z}$ and $x_2 \in \mathbb{R}$.

Exercise 1.9

Let $\mathbf{c} \in \mathbb{R}^m$ be fixed. Show that the constant function $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ given by

$$f(\mathbf{x}) = \mathbf{c}$$

is continuous on $\mathbb{R}^n$.

In the next worked exercise we apply the definition of continuity to the distance function $d^{(n)}(\mathbf{x}, \mathbf{u})$ that maps a point in $\mathbb{R}^n$ to its distance from a fixed point $\mathbf{u}$; it is thus a function from $\mathbb{R}^n$ to $\mathbb{R}$.

Worked Exercise 1.15

Let $\mathbf{u}$ be a fixed point in $\mathbb{R}^n$. Prove that the function $g: \mathbb{R}^n \rightarrow \mathbb{R}$ given by

$$g(\mathbf{x}) = d^{(n)}(\mathbf{x}, \mathbf{u})$$

is continuous on $\mathbb{R}^n$. (The graph of this function is shown in Figure 1.2.)

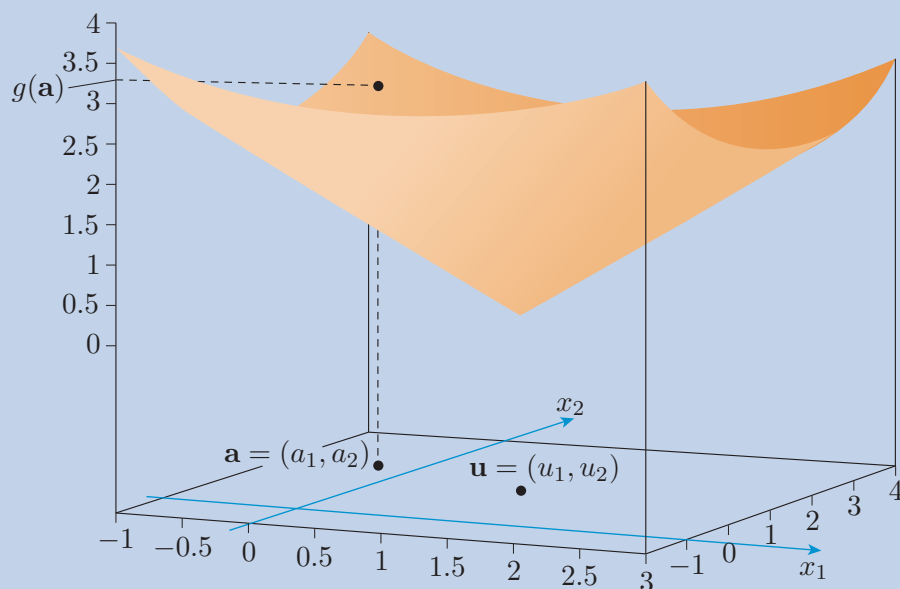


Figure 1.2 Graph of function for Worked Exercise 1.15

Solution

Let $\mathbf{a}$ be a general point of $\mathbb{R}^n$ and let $(\mathbf{x}_k)$ be a sequence in $\mathbb{R}^n$ such that $\mathbf{x}_k \rightarrow \mathbf{a}$. We need to show that $(d^{(1)}(g(\mathbf{x}_k), g(\mathbf{a})))$ is a null sequence in $\mathbb{R}$.

We find an upper bound for $d^{(1)}(g(\mathbf{x}_k), g(\mathbf{a})) = |g(\mathbf{x}_k) - g(\mathbf{a})|$ in terms of $d^{(n)}(\mathbf{x}_k, \mathbf{a})$. By the definition of g ,

$$|g(\mathbf{x}_k) - g(\mathbf{a})| = |d^{(n)}(\mathbf{x}_k, \mathbf{u}) - d^{(n)}(\mathbf{a}, \mathbf{u})|.$$

From the Reverse Triangle Inequality, Proposition 1.4, we see that

$$|d^{(n)}(\mathbf{x}_k, \mathbf{u}) - d^{(n)}(\mathbf{a}, \mathbf{u})| = |d^{(n)}(\mathbf{u}, \mathbf{x}_k) - d^{(n)}(\mathbf{u}, \mathbf{a})| \leq d^{(n)}(\mathbf{x}_k, \mathbf{a}).$$

So, $0 \leq |g(\mathbf{x}_k) - g(\mathbf{a})| \leq d^{(n)}(\mathbf{x}_k, \mathbf{a})$.

But $\mathbf{x}_k \rightarrow \mathbf{a}$, and so $(d^{(n)}(\mathbf{x}_k, \mathbf{a}))$ is a (real) null sequence. Hence, the Squeeze Rule implies that $g(\mathbf{x}_k) \rightarrow g(\mathbf{a})$.

Thus g is continuous at $\mathbf{a}$. Since $\mathbf{a}$ is an arbitrary point of $\mathbb{R}^n$, g is continuous on $\mathbb{R}^n$.

In this solution we need to keep a clear head, since the distance function is used in the domain and also as the function in the codomain.

As with the functions discussed in the previous chapter, there are combination rules for continuous functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ and these often make it easier to show that a given function is continuous.

Combination rules for continuous functions from $\mathbb{R}^n$ to $\mathbb{R}^m$

Let $A \subseteq \mathbb{R}^n$, and let $f: A \rightarrow \mathbb{R}^m$ and $g: A \rightarrow \mathbb{R}^m$ be continuous on A .

Then the following functions are continuous on A :

Sum Rule $f + g: A \rightarrow \mathbb{R}^m$, defined by $(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$

Multiple Rule $\lambda f: A \rightarrow \mathbb{R}^m$ where $\lambda \in \mathbb{R}$, defined by $(\lambda f)(\mathbf{x}) = \lambda \times f(\mathbf{x})$.

These rules can be quickly deduced from the combination rules for convergent sequences in higher dimensions.

There is also a Restriction Rule for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ that allows us to restrict the domain of a continuous function to a smaller set and still have a continuous function.

Proposition 1.16 Restriction Rule

Let $A \subseteq \mathbb{R}^n$, let $f: A \rightarrow \mathbb{R}^m$ be continuous on A and let $B \subseteq A$. Then the restricted function $f|_B: B \rightarrow \mathbb{R}^m$ defined by $f|_B(\mathbf{x}) = f(\mathbf{x})$ for $\mathbf{x} \in B$ is continuous on B .

We can also generalise the Composition Rule. If the first function is from $\mathbb{R}^n$ to $\mathbb{R}^m$, then the second function must have its domain in $\mathbb{R}^m$, but may have any dimension for its codomain. Again we state it without proof.

Proposition 1.17 *Composition Rule*

Let $A \subseteq \mathbb{R}^n$ and let $f: A \rightarrow \mathbb{R}^m$ be continuous on A . Let $f(A) \subseteq B \subseteq \mathbb{R}^m$ and let $g: B \rightarrow \mathbb{R}^k$ be continuous on B . Then the composed function $g \circ f: A \rightarrow \mathbb{R}^k$ is continuous on A .

Exercise 1.10

Without using sequences, prove that the function $g: \mathbb{R}^n \rightarrow \mathbb{R}$ given by

$$g(\mathbf{x}) = d^{(n)}(\mathbf{x}, \mathbf{0})^2 \quad \left(= d^{(n)}(\mathbf{x}, \mathbf{0}) \times d^{(n)}(\mathbf{x}, \mathbf{0}) \right)$$

is continuous on $\mathbb{R}^n$.

This function maps $\mathbf{x}$ to the square of its distance from $\mathbf{0}$.

See Chapter 13, Definition 3.10.

The Composition Rule for continuous functions is particularly useful when used in combination with projection functions from $\mathbb{R}^n$ to $\mathbb{R}$, which are similar to the projection functions from $\mathbb{R}^2$ to $\mathbb{R}$ introduced in Chapter 13.

Definition 1.18 *Projection functions on $\mathbb{R}^n$*

The n functions $p_j: \mathbb{R}^n \rightarrow \mathbb{R}$ given by

$$p_j(x_1, x_2, \dots, x_n) = x_j, \quad \text{for } j = 1, 2, \dots, n,$$

are known as **projection functions**.

Each of these projection functions maps a point $\mathbf{x} = (x_1, x_2, \dots, x_n)$ onto one of its coordinates. Thus $p_1(\mathbf{x}) = x_1$, $p_2(\mathbf{x}) = x_2$, and so on.

Exercise 1.11

Show that the projection functions are continuous on $\mathbb{R}^n$.

Note that the effect of composing each separate p_j with a function f is to produce the individual component function values of f in the codomain. For example, composing p_1 , p_2 and p_3 with the function

$$h: \mathbb{R}^5 \rightarrow \mathbb{R}^3, \quad h(x_1, x_2, x_3, x_4, x_5) = (2x_1, x_2^2 + 3x_4, x_3 + x_5),$$

gives

$$\begin{aligned} (p_1 \circ h)(x_1, x_2, x_3, x_4, x_5) &= 2x_1, \\ (p_2 \circ h)(x_1, x_2, x_3, x_4, x_5) &= x_2^2 + 3x_4, \\ (p_3 \circ h)(x_1, x_2, x_3, x_4, x_5) &= x_3 + x_5. \end{aligned}$$

The Composition Rule and the projection functions $p_j: \mathbb{R}^m \rightarrow \mathbb{R}$ give us a useful technique for proving that a function f from $\mathbb{R}^n$ to $\mathbb{R}^m$ is continuous. The main idea is that, instead of having to prove directly that f is continuous, we form the composition of each of the projection functions p_j with the function f , and then prove that each of these functions $p_j \circ f$ is continuous. Since the functions $p_j \circ f$ are functions from $\mathbb{R}^n$ to $\mathbb{R}$, this is usually simpler than proving directly that f is continuous. The following result enables us to use projection functions in this way.

Proposition 1.19

Let $A \subseteq \mathbb{R}^n$ and let $f: A \rightarrow \mathbb{R}^m$.

The function f is continuous at $\mathbf{a} \in A$ if, and only if, for each $j = 1, 2, \dots, m$, the function $p_j \circ f: A \rightarrow \mathbb{R}$ is continuous at $\mathbf{a}$, where $p_j: \mathbb{R}^m \rightarrow \mathbb{R}$ is the j th projection function.

Proof The proof is in two parts. In each part we deal with the m functions $p_1 \circ f, p_2 \circ f, \dots, p_m \circ f$ simultaneously.

First, we prove that if f is continuous at $\mathbf{a}$ then the functions $p_j \circ f$ are also continuous at $\mathbf{a}$. This follows directly from the Composition Rule for continuous functions. Since f is continuous at $\mathbf{a}$ and each of the p_j are continuous on $\mathbb{R}^m$, the functions $p_j \circ f$ are also continuous at $\mathbf{a}$, for $j = 1, 2, \dots, m$.

Conversely, we now prove that if the functions $p_j \circ f$ are continuous at $\mathbf{a}$ then f is also continuous at $\mathbf{a}$. To do this, let $(\mathbf{x}_k)$ be a sequence in A that converges to $\mathbf{a}$. Then, since the functions $p_j \circ f$ are continuous at $\mathbf{a}$, we have that

$$(p_j \circ f)(\mathbf{x}_k) \rightarrow (p_j \circ f)(\mathbf{a}) \text{ as } k \rightarrow \infty \text{ for each } j = 1, 2, \dots, m.$$

The value of f at any point $\mathbf{x} \in A$ can be expressed as

$$f(\mathbf{x}) = ((p_1 \circ f)(\mathbf{x}), \dots, (p_m \circ f)(\mathbf{x})).$$

Using Proposition 1.9, it follows that

$$\begin{aligned} f(\mathbf{x}_k) &= ((p_1 \circ f)(\mathbf{x}_k), \dots, (p_m \circ f)(\mathbf{x}_k)) \\ &\rightarrow ((p_1 \circ f)(\mathbf{a}), \dots, (p_m \circ f)(\mathbf{a})) = f(\mathbf{a}) \text{ as } k \rightarrow \infty. \end{aligned}$$

Thus f is continuous at $\mathbf{a}$. ■

This is the useful part of Proposition 1.19.

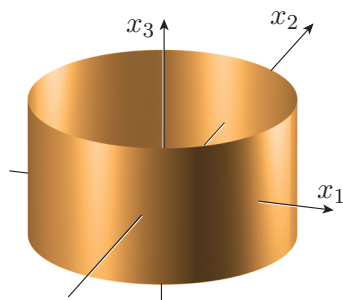


Figure 1.3 The image set $f(A)$ in Worked Exercise 1.20

See Worked Exercise 3.14 in Chapter 13.

Worked Exercise 1.20

Let $A = [0, 1] \times [-1, 1] = \{(x_1, x_2) : 0 \leq x_1 \leq 1 \text{ and } -1 \leq x_2 \leq 1\}$ and define $f: A \rightarrow \mathbb{R}^3$ by

$$f(x_1, x_2) = (\cos 2\pi x_1, \sin 2\pi x_1, x_2).$$

Prove that f is continuous on A . (Note that the image set $f(A)$ is a cylinder, as shown in Figure 1.3.)

Solution

We consider the composition of f with the three projection functions p_1, p_2 and p_3 from $\mathbb{R}^3$ to $\mathbb{R}$, given by

$$(p_1 \circ f)(x_1, x_2) = \cos 2\pi x_1,$$

$$(p_2 \circ f)(x_1, x_2) = \sin 2\pi x_1,$$

$$(p_3 \circ f)(x_1, x_2) = x_2.$$

If we can show that each $p_j \circ f$ is continuous on A , then it follows from Proposition 1.19 that f is continuous on A .

To show that $(p_1 \circ f)(x_1, x_2) = \cos 2\pi x_1$ is continuous on A , we show that the constituent parts are continuous and then combine them using the combination rules.

In the previous chapter, we proved the continuity of the function $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ given by $g(\mathbf{x}) = u_1 x_1 + u_2 x_2$, where u_1 and u_2 are fixed. It follows, on taking $u_1 = 2\pi$ and $u_2 = 0$, that the function given by $(x_1, x_2) \mapsto 2\pi x_1$ is continuous.

We are also restricting the functions to the set A , and so we need to apply the Restriction Rule for functions from $\mathbb{R}^2$ to $\mathbb{R}$. Then, since $x \mapsto \cos x$ is a basic continuous function from $\mathbb{R}$ to $\mathbb{R}$, we can apply the Composition Rule for continuous functions. This shows that $(x_1, x_2) \mapsto \cos 2\pi x_1$ is continuous.

The proof of continuity for $p_2 \circ f$ is similar, except that we use the basic continuous function $x \mapsto \sin x$.

Finally, we note that $p_3 \circ f$ is a function of the form $g(\mathbf{x}) = u_1 x_1 + u_2 x_2$, as considered above, with $\mathbf{u} = (0, 1)$ and restricted to the set A . Hence, by the Restriction Rule, it is continuous. (Alternatively we could simply note that $p_3 \circ f$ is the projection function $p_2: \mathbb{R}^2 \rightarrow \mathbb{R}$ and hence is continuous.)

Since all three compositions of f with a projection function are continuous on A , we conclude from Proposition 1.19 that f is continuous on A .

Exercise 1.12

Let $\mathbf{v} = (v_1, v_2, v_3)$ be a fixed point in $\mathbb{R}^3$. Define $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ by

$$f(x_1, x_2, x_3) = (v_3x_2 - v_2x_3, v_1x_3 - v_3x_1, v_2x_1 - v_1x_2).$$

Prove that f is continuous on $\mathbb{R}^3$.

1.5 Proof of the Triangle Inequality

In this subsection we give a proof of the Triangle Inequality

This subsection is not assessed.

$$d^{(n)}(\mathbf{a}, \mathbf{c}) \leq d^{(n)}(\mathbf{a}, \mathbf{b}) + d^{(n)}(\mathbf{b}, \mathbf{c})$$

using only algebraic methods. The proof makes use of another inequality, the *Cauchy–Schwarz Inequality*, which has important applications in many areas of mathematics. We first state and prove this inequality.

Theorem 1.21 *Cauchy–Schwarz Inequality*

Let $(r_1, r_2, \dots, r_n)$ and $(s_1, s_2, \dots, s_n)$ be points in $\mathbb{R}^n$. Then

$$\left(\sum_{j=1}^n r_j s_j \right)^2 \leq \sum_{j=1}^n r_j^2 \sum_{j=1}^n s_j^2, \quad (4)$$

with equality if, and only if, there is $\lambda \in \mathbb{R}$ for which $s_j = \lambda r_j$ for each j .

A useful technique for understanding a general statement is to specialise to small particular values.

When $n = 1$, the right-hand side of the Cauchy–Schwarz Inequality is $(r_1^2)(s_1^2)$ and the left-hand side is $(r_1 s_1)^2$. So taking the left-hand side from the right-hand side gives

$$(r_1^2)(s_1^2) - (r_1 s_1)^2 = 0,$$

and so inequality (4) certainly holds for $n = 1$.

When $n = 2$, the right-hand side becomes

$$(r_1^2 + r_2^2)(s_1^2 + s_2^2)$$

and the left-hand side is

$$(r_1 s_1 + r_2 s_2)^2.$$

Thus subtracting the left-hand side from the right-hand side gives

$$(r_1^2 + r_2^2)(s_1^2 + s_2^2) - (r_1 s_1 + r_2 s_2)^2$$

and, multiplying out, we obtain

$$\begin{aligned} & (r_1^2 s_1^2 + r_1^2 s_2^2 + r_2^2 s_1^2 + r_2^2 s_2^2) - (r_1^2 s_1^2 + 2r_1 r_2 s_1 s_2 + r_2^2 s_2^2) \\ &= r_1^2 s_2^2 + r_2^2 s_1^2 - 2r_1 r_2 s_1 s_2 \\ &= (r_1 s_2 - r_2 s_1)^2 \geq 0. \end{aligned}$$

So (4) holds (with equality if, and only if, there is $\lambda \in \mathbb{R}$ for which $s_1 = \lambda r_1$ and $s_2 = \lambda r_2$).

We could give a proof of the Cauchy–Schwarz Inequality by developing these ideas for a general value of n . However, we give a different proof, which involves a strategy based on some properties of quadratic equations with which you may well be familiar.

Let $A, B, C \in \mathbb{R}$, where $A \neq 0$, and suppose that

$$Ax^2 + Bx + C \geq 0, \quad \text{for each } x \in \mathbb{R}.$$

Then the quadratic equation $Ax^2 + Bx + C = 0$ has at most one real root. For, if the equation had two distinct real roots, then the graph of $f(x) = Ax^2 + Bx + C$ would cut the x -axis in two places and there would be a range of x for which $f(x) < 0$, i.e. $Ax^2 + Bx + C < 0$.

The roots of the quadratic equation are given by

$$x = \frac{-B \pm \sqrt{B^2 - 4AC}}{2A}.$$

When there is at most one real root, it must be the case that $B^2 - 4AC \leq 0$, so that $B^2 \leq 4AC$. In particular, there is exactly one real root in the case that $B^2 = 4AC$; otherwise there are none. This is the idea we use.

Proof of the Cauchy–Schwarz Inequality If $r_j = 0$ for $j = 1, 2, \dots, n$, then inequality (4) simplifies to become

$$0 \leq 0 \times \sum_{j=1}^n s_j^2 = 0,$$

which is clearly true. Therefore, we need only consider cases where $(r_1, r_2, \dots, r_n) \neq (0, 0, \dots, 0)$.

Let $x \in \mathbb{R}$ and consider the expression

$$\sum_{j=1}^n (r_j x + s_j)^2.$$

Since a sum of squares is always non-negative, it follows that for any real number x ,

$$\sum_{j=1}^n (r_j x + s_j)^2 \geq 0,$$

with equality if, and only if, each term in the sum is 0 (which means that $s_j = -xr_j$ for each j).

Multiplying out the left-hand side of this inequality, we obtain

$$\sum_{j=1}^n (r_j^2 x^2 + 2r_j s_j x + s_j^2) \geq 0,$$

or equivalently

$$\left(\sum_{j=1}^n r_j^2 \right) x^2 + 2 \left(\sum_{j=1}^n r_j s_j \right) x + \left(\sum_{j=1}^n s_j^2 \right) \geq 0.$$

The left-hand side is a quadratic expression $Ax^2 + Bx + C$ with

$$A = \sum_{j=1}^n r_j^2, \quad B = 2 \sum_{j=1}^n r_j s_j \quad \text{and} \quad C = \sum_{j=1}^n s_j^2.$$

A is not zero since $(r_1, r_2, \dots, r_n) \neq (0, 0, \dots, 0)$.

So we have, for each choice of $x \in \mathbb{R}$,

$$Ax^2 + Bx + C \geq 0,$$

with equality if, and only if, $s_j = -xr_j$ for each j . Hence

$$B^2 \leq 4AC.$$

Substituting for A , B and C in this inequality, we obtain

$$4 \left(\sum_{j=1}^n r_j s_j \right)^2 \leq 4 \left(\sum_{j=1}^n r_j^2 \right) \left(\sum_{j=1}^n s_j^2 \right).$$

Dividing each side by 4 gives the required inequality. ■

If we take the square root of both sides of equation (4), we obtain the inequality

$$\sum_{j=1}^n r_j s_j \leq \left(\sum_{j=1}^n r_j^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^n s_j^2 \right)^{\frac{1}{2}}. \quad (5)$$

This inequality is also known as the Cauchy–Schwarz Inequality.

We can now prove the Triangle Inequality (M3) for the Euclidean distance on $\mathbb{R}^n$.

Triangle Inequality

For each $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$, $d^{(n)}(\mathbf{a}, \mathbf{c}) \leq d^{(n)}(\mathbf{a}, \mathbf{b}) + d^{(n)}(\mathbf{b}, \mathbf{c})$.

In the proof, we write the Triangle Inequality in a form that enables us to see that it follows from the Cauchy–Schwarz Inequality. To do this we use the formula for the Euclidean distance function. There is nothing very deep here, but the algebraic manipulation requires some care, as it involves sums of n real numbers and their squares.

Proof Let $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ be in $\mathbb{R}^n$. For convenience, we introduce the abbreviations $\mathbf{r} = \mathbf{a} - \mathbf{b}$ and $\mathbf{s} = \mathbf{b} - \mathbf{c}$, so that $\mathbf{r} + \mathbf{s} = \mathbf{a} - \mathbf{c}$. Thus

$$\begin{aligned} d^{(n)}(\mathbf{a}, \mathbf{b}) &= d^{(n)}(\mathbf{a} - \mathbf{b}, \mathbf{0}) = d^{(n)}(\mathbf{r}, \mathbf{0}) = \left(\sum_{j=1}^n r_j^2 \right)^{\frac{1}{2}}, \\ d^{(n)}(\mathbf{b}, \mathbf{c}) &= d^{(n)}(\mathbf{b} - \mathbf{c}, \mathbf{0}) = d^{(n)}(\mathbf{s}, \mathbf{0}) = \left(\sum_{j=1}^n s_j^2 \right)^{\frac{1}{2}}, \\ d^{(n)}(\mathbf{a}, \mathbf{c}) &= d^{(n)}(\mathbf{a} - \mathbf{c}, \mathbf{0}) = d^{(n)}(\mathbf{r} + \mathbf{s}, \mathbf{0}) = \left(\sum_{j=1}^n (r_j + s_j)^2 \right)^{\frac{1}{2}}. \end{aligned}$$

Squaring and expanding, we find that

$$\begin{aligned} \left(d^{(n)}(\mathbf{a}, \mathbf{c}) \right)^2 &= \sum_{j=1}^n (r_j + s_j)^2 \\ &= \left(\sum_{j=1}^n r_j^2 \right) + 2 \left(\sum_{j=1}^n r_j s_j \right) + \left(\sum_{j=1}^n s_j^2 \right). \end{aligned}$$

We are aiming to obtain an expression to which we can apply the Cauchy–Schwarz Inequality.

We square to avoid the complication of having to deal with the square root.

If we now use the Cauchy–Schwarz Inequality, equation (5), to find an upper bound for the middle term on the right-hand side, we obtain

$$\begin{aligned} \left(d^{(n)}(\mathbf{a}, \mathbf{c}) \right)^2 &\leq \left(\sum_{j=1}^n r_j^2 \right) + 2 \left(\sum_{j=1}^n r_j^2 \right)^{\frac{1}{2}} \left(\sum_{j=1}^n s_j^2 \right)^{\frac{1}{2}} + \left(\sum_{j=1}^n s_j^2 \right) \\ &= \left(\left(\sum_{j=1}^n r_j^2 \right)^{\frac{1}{2}} + \left(\sum_{j=1}^n s_j^2 \right)^{\frac{1}{2}} \right)^2 \\ &= \left(d^{(n)}(\mathbf{a}, \mathbf{b}) + d^{(n)}(\mathbf{b}, \mathbf{c}) \right)^2. \end{aligned}$$

Taking square roots, we obtain

$$d^{(n)}(\mathbf{a}, \mathbf{c}) \leq d^{(n)}(\mathbf{a}, \mathbf{b}) + d^{(n)}(\mathbf{b}, \mathbf{c}),$$

which completes the proof of the Triangle Inequality for the usual Euclidean distance in $\mathbb{R}^n$. ■

2 Introducing metric spaces

We now introduce the idea of a metric space, and show how this concept allows us to generalise the notion of continuity. We will then concentrate on looking at some examples of metric spaces and defer further discussion of continuous functions between metric spaces to the next chapter.

Subsection 2.1 shows how properties of the Euclidean distance function on $\mathbb{R}^n$ can be generalised to define a metric on any set to create a metric space.

You should not expect to have a firm grasp of the idea of a metric space by the end of this section: this will come as you see more examples of metric spaces in Chapter 15 and later chapters.

2.1 The definition of a metric space

Looking back at the definition of continuity for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ that we gave in Section 1, we see that it depends on:

- the Euclidean distance in $\mathbb{R}^n$, $d^{(n)}$
- the Euclidean distance in $\mathbb{R}^m$, $d^{(m)}$
- the definition of convergence for sequences in both $\mathbb{R}^n$ and $\mathbb{R}^m$.

Recall that if $\mathbf{a}, \mathbf{b} \in \mathbb{R}^k$, then $d^{(k)}(\mathbf{a}, \mathbf{b})$ is given by $\sqrt{\sum_{i=1}^k (b_i - a_i)^2}$.

Moreover, the definition of convergence for sequences also depends on the notion of distance. Thus a notion of distance (in both the domain and the codomain) is a fundamental part of our definition of continuity. Hence to define a notion of continuity for functions between arbitrary sets X and Y , we must first have effective notions of distance appropriate to the sets X and Y . An obvious strategy to do this is to try to find some properties of the Euclidean distance that we would expect any sensible notion of distance to possess.

In Proposition 1.3, we showed that the Euclidean distance function has three properties.

The *Euclidean distance function* $d^{(n)}: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ has the following properties.

For each $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^n$:

- (M1) $d^{(n)}(\mathbf{a}, \mathbf{b}) \geq 0$, with equality holding if, and only if, $\mathbf{a} = \mathbf{b}$
- (M2) $d^{(n)}(\mathbf{a}, \mathbf{b}) = d^{(n)}(\mathbf{b}, \mathbf{a})$
- (M3) $d^{(n)}(\mathbf{a}, \mathbf{c}) \leq d^{(n)}(\mathbf{a}, \mathbf{b}) + d^{(n)}(\mathbf{b}, \mathbf{c})$ (the *Triangle Inequality*).

Properties (M1)–(M3) do not use any special properties of the set $\mathbb{R}^n$, so it seems reasonable to say that for any set X , a function $d: X \times X \rightarrow \mathbb{R}$ is a *distance function* if it satisfies them. It turns out that these are indeed the appropriate properties of distance on which to base our definition. We call such a distance function a *metric*, and we call a set with such a function defined on it a *metric space*.

The word *metric* comes from the Greek word *μετρῶν* (*metron*), meaning ‘measure’, an instrument for measuring.

Definition 2.1 Metric

Let X be a set. A **metric** on X is a function $d: X \times X \rightarrow \mathbb{R}$ that satisfies the following three conditions.

For each $a, b, c \in X$:

- (M1) $d(a, b) \geq 0$, with equality holding if, and only if, $a = b$
- (M2) $d(a, b) = d(b, a)$
- (M3) $d(a, c) \leq d(a, b) + d(b, c)$ (the **Triangle Inequality**).

The set X , together with a metric d on X , is called a **metric space**, and is denoted by (X, d) .

This definition was proposed by Maurice Fréchet, whose work is discussed in the *History Reader*.

A ‘point’ may be nothing like a dot in the plane: X could be a set of functions and then a point in X would be a function.

Remarks

1. Mathematicians usually refer to the members of the set X as **points**, to emphasise the analogy with the points of a line, a plane or three-dimensional space. Similarly, mathematicians usually refer to $d(a, b)$ as the **distance** between a and b . The theory of metric spaces is the study of those properties of sets of points that depend only on distance.
2. Condition (M1) says that distance is a non-negative quantity, and that the only point of a metric space that is at zero distance from a given point is that point itself.
3. Condition (M2) says that the distance from a point a to a point b is precisely the same as the distance from b to a . (The metric is *symmetric*. It doesn’t matter whether you are going from a to b or from b to a – the distance is the same.)
4. Condition (M3) tells us that $d(a, c)$ gives the ‘shortest’ distance between a and c . For if we go directly from a to c , that gives a distance of $d(a, c)$. If we make a detour to b along the way, we must go a distance $d(a, b)$ to get to b , and then an additional distance $d(b, c)$ to get from b to c (Figure 2.1). Condition (M3) tells us that the total distance travelled must then be at least as great as the ‘direct’ distance $d(a, c)$.
5. Since our general definition is modelled on the corresponding properties of the Euclidean distance function $d^{(n)}$, it follows that $(\mathbb{R}^n, d^{(n)})$ is a metric space, for each $n \in \mathbb{N}$. It is known as **Euclidean n -space**. Furthermore, in the context of metric spaces, the Euclidean distance function $d^{(n)}$ is often referred to as the **Euclidean metric** for $\mathbb{R}^n$. These are our first examples of metric spaces.

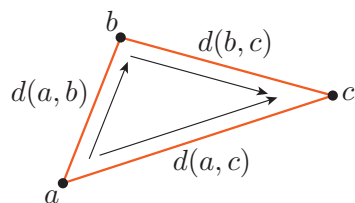


Figure 2.1 Shortest distance between a and c

If we look back at the proof of the Reverse Triangle Inequality for the Euclidean metric on $\mathbb{R}^n$ (Proposition 1.4), we see that the proof makes use only of properties (M1)–(M3) of the Euclidean metric. So it is no surprise to learn that there is a Reverse Triangle Inequality for any metric space.

Proposition 2.2 *Reverse Triangle Inequality for metric spaces*

Let (X, d) be a metric space. For each $a, b, c \in X$,
 (M3a) $d(b, c) \geq |d(a, c) - d(a, b)|$.

The proof is essentially the same as that for Proposition 1.4, and so we omit it.

2.2 Examples of metric spaces

We now give a few examples of metric spaces with the aim of getting you used to showing that a distance function is a metric. The examples we consider here are not very sophisticated and may not persuade you of the utility of this new concept. However, in the next chapter we investigate ways of defining distances between more interesting ‘points’, such as between pairs of functions or pairs of sequences – this is when we will start to see the real power of metrics.

The discrete metric

Our first example is simple but important, since it shows that it is possible to define a metric on any set X .

Definition 2.3 *Discrete metric*

Let X be a set. The **discrete metric** on X is the function $d_0: X \times X \rightarrow \mathbb{R}$ defined by

$$d_0(a, b) = \begin{cases} 0, & \text{if } a = b, \\ 1, & \text{if } a \neq b. \end{cases}$$

We show that d_0 is a metric after the remarks.

Remarks

1. When $X = \emptyset$, the empty set, the definition still makes sense but is not very interesting.
2. The d_0 -distance between any two *distinct* points of a set X is always equal to 1. In particular, if we consider the case when $X = \mathbb{R}$, the real line, we find that d_0 is a different metric from the usual Euclidean distance. (To see this, observe that $d_0(0, 2) = 1$ whereas $d^{(1)}(0, 2) = 2$.) Thus the discrete metric gives us a second notion of distance on $\mathbb{R}$. In fact, if (X, d) is any metric space whose metric d is not d_0 , then (X, d_0) is a second metric space with underlying set X .

The discrete metric is about the simplest possible notion of distance there can be and acts to ‘isolate’ points of a space: they are all at distance 1 from one another.

Proposition 2.4

Let X be a set. Then (X, d_0) is a metric space.

Proof Let X be a set. In order to show that d_0 is a metric on X , we must show that d_0 satisfies conditions (M1)–(M3).

If $X = \emptyset$, then X has no elements and so each of (M1)–(M3) vacuously holds: there is nothing to prove. So suppose instead that X is non-empty.

(M1) Since d_0 can take only the values 0 and 1, we have $d_0(a, b) \geq 0$ for each choice of $a, b \in X$.

The definition of d_0 implies immediately that $d_0(a, b) = 0$ if, and only if, $a = b$.

Thus d_0 satisfies (M1).

(M2) If $a = b$, then $d_0(a, b) = 0 = d_0(b, a)$.

If $a \neq b$, then $d_0(a, b) = 1 = d_0(b, a)$.

Thus, for each $a, b \in X$, (M2) holds.

(M3) Let $a, b, c \in X$. We examine the two possible cases: $d_0(a, c) = 0$ and $d_0(a, c) = 1$.

Suppose $d_0(a, c) = 0$ (so $a = c$). Since $d_0(a, b)$ and $d_0(b, c)$ are non-negative, it follows that

$$d_0(a, b) + d_0(b, c) \geq 0 = d_0(a, c).$$

Now suppose $d_0(a, c) = 1$; then $a \neq c$ and so b cannot equal both a and c . Hence, from (M1), at least one of $d_0(a, b)$ and $d_0(b, c)$ is non-zero and so must equal 1. Thus

$$d_0(a, b) + d_0(b, c) \geq 1 = d_0(a, c).$$

Hence, in both cases, (M3) holds.

Since d_0 satisfies (M1)–(M3), it is a metric on X and (X, d_0) is a metric space. ■

At first appearance, the simplicity of the definition of the discrete metric d_0 may make it seem unimportant, but this is not the case. Its utility lies in the fact that it gives an ‘extreme’ example of a metric space – no other definition of distance so completely ignores any structure that may be present in the underlying set X .

We use the discrete metric for a number of purposes, principally to test properties that we suspect may hold for all metric spaces. If such a property fails for (X, d_0) , then we know that our suspicion was false.

Exercise 2.1

Let $X = \{x, y, z\}$ and define $d: X \times X \rightarrow \mathbb{R}$ by

$$\begin{aligned} d(x, x) &= d(y, y) = d(z, z) = 0, \\ d(x, y) &= d(y, x) = 1, \\ d(y, z) &= d(z, y) = 2, \\ d(x, z) &= d(z, x) = 4. \end{aligned}$$

Determine whether d is a metric on X .

The taxicab metric: an alternative metric for the plane

We have already observed that the usual definition of distance on the plane defines a metric and we have just seen that it is possible to define a very simple notion of distance on the plane, namely the discrete metric given previously. Here we introduce another natural way to measure distance between points in the plane, the taxicab metric.

Recall that the usual definition of distance on the plane is given by

$$d^{(2)}(\mathbf{a}, \mathbf{b}) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2} \quad \text{for } \mathbf{a}, \mathbf{b} \in \mathbb{R}^2.$$

What happens if we change the formula we use for distance on the right-hand side of this expression to something that is perhaps easier to calculate? For example, consider the function $e_1: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ defined by

$$e_1(\mathbf{a}, \mathbf{b}) = |b_1 - a_1| + |b_2 - a_2| \quad \text{for } \mathbf{a}, \mathbf{b} \in \mathbb{R}^2.$$

As we will shortly see, this *does* define an alternative metric on the plane. First let us understand how this distance function behaves.

Exercise 2.2

Find:

- (a) $e_1((0, 0), (1, 0))$; (b) $e_1((0, 0), (0, 1))$;
(c) $e_1((0, 1), (1, 0))$.

You may have realised from your work on Exercise 2.2 that one way to understand $e_1(\mathbf{a}, \mathbf{b})$ is to draw a right-angled triangle with $\mathbf{a}$ and $\mathbf{b}$ at its two non-right-angled vertices, as shown in Figure 2.2. Then $e_1(\mathbf{a}, \mathbf{b})$ is the sum of the distances along the sides parallel to the axes. Another way to interpret it is to consider $\mathbf{a}$ and $\mathbf{b}$ as representing intersections in a city where the roads form a rectangular grid, as Figure 2.3 illustrates. Then $e_1(\mathbf{a}, \mathbf{b})$ is the shortest distance that can be travelled by a vehicle, such as a taxi, to get from $\mathbf{a}$ to $\mathbf{b}$. This interpretation leads to the common name *taxicab metric* for e_1 .

The reason for the subscript 1 in the notation e_1 will become clear in Chapter 15.

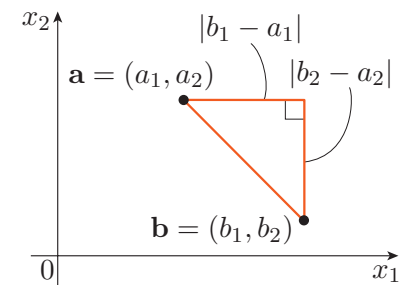


Figure 2.2 Visual representation of $e_1(\mathbf{a}, \mathbf{b})$

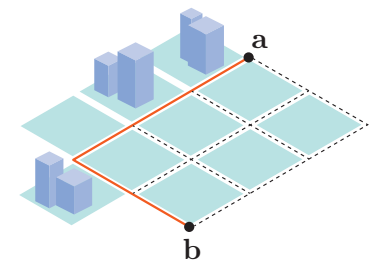


Figure 2.3 Alternative interpretation of $e_1(\mathbf{a}, \mathbf{b})$

Let us verify that e_1 does define a metric: we must check that it satisfies conditions (M1)–(M3).

- (M1) For each $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in \mathbb{R}^2$, $|b_1 - a_1|$ and $|b_2 - a_2|$ are non-negative, and hence so is their sum: thus $e_1(\mathbf{a}, \mathbf{b}) \geq 0$.

For each $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$,

$$e_1(\mathbf{a}, \mathbf{a}) = |a_1 - a_1| + |a_2 - a_2| = 0 + 0 = 0.$$

Conversely, suppose $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in \mathbb{R}^2$ are such that $e_1(\mathbf{a}, \mathbf{b}) = 0$. Then

$$0 = |b_1 - a_1| + |b_2 - a_2|,$$

which implies that both $|b_1 - a_1| = 0$ and $|b_2 - a_2| = 0$. Thus $a_1 = b_1$ and $a_2 = b_2$; that is, $\mathbf{a} = \mathbf{b}$.

This proves that condition (M1) is satisfied.

- (M2) Let $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in \mathbb{R}^2$. Using that the modulus function has property (M2), we obtain

$$\begin{aligned} e_1(\mathbf{a}, \mathbf{b}) &= |b_1 - a_1| + |b_2 - a_2| \\ &= |a_1 - b_1| + |a_2 - b_2| = e_1(\mathbf{b}, \mathbf{a}). \end{aligned}$$

This proves that condition (M2) is satisfied by e_1 .

- (M3) Let $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2), \mathbf{c} = (c_1, c_2) \in \mathbb{R}^2$. Then, using that the modulus function has property (M3),

$$\begin{aligned} e_1(\mathbf{a}, \mathbf{c}) &= |c_1 - a_1| + |c_2 - a_2| \\ &\leq (|c_1 - b_1| + |b_1 - a_1|) + (|c_2 - b_2| + |b_2 - a_2|) \\ &= (|b_1 - a_1| + |b_2 - a_2|) + (|c_1 - b_1| + |c_2 - b_2|) \\ &= e_1(\mathbf{a}, \mathbf{b}) + e_1(\mathbf{b}, \mathbf{c}). \end{aligned}$$

This proves that condition (M3) is satisfied by e_1 .

We conclude that e_1 is a metric on the plane.

Exercise 2.3

Determine which of the following functions define a metric on $\mathbb{R}$:

- (a) $d(a, b) = \sqrt{b^2 - a^2}$; (b) $d(a, b) = \sqrt{|b|^3 + |a|^3}$;
(c) $d(a, b) = \sqrt[3]{|b^3 - a^3|}$.

The modulus function is the Euclidean distance function on $\mathbb{R}$, and so is a metric.

A ‘mixed’ metric in the plane

Our next example of a metric is a mixture of the discrete metric and the modulus function.

For points $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2)$ in the plane, define the function $d: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ by

$$d(\mathbf{a}, \mathbf{b}) = |b_1 - a_1| + d_0(a_2, b_2),$$

where d_0 denotes the discrete metric on $\mathbb{R}$.

Our goal is to show that this is indeed a metric on the plane. To do this, we make extensive use of the fact that the modulus function and the discrete metric both satisfy properties (M1)–(M3) on $\mathbb{R}$.

Exercise 2.4

Show that d has property (M1).

Exercise 2.5

Show that d has property (M2).

We verify the Triangle Inequality as a worked exercise.

Worked Exercise 2.5

Show that d has property (M3).

Solution

We must show that for any choice of $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^2$, we have

$$d(\mathbf{a}, \mathbf{c}) \leq d(\mathbf{a}, \mathbf{b}) + d(\mathbf{b}, \mathbf{c}).$$

As with the previous exercises, we use the fact that both $|\cdot|$ and d_0 have property (M3) when viewed as metrics on $\mathbb{R}$.

Let $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2)$, $\mathbf{c} = (c_1, c_2) \in \mathbb{R}^2$. Then

$$\begin{aligned} d(\mathbf{a}, \mathbf{c}) &= |a_1 - c_1| + d_0(c_2, a_2), \text{ by the definition of } d \\ &\leq |a_1 - b_1| + |b_1 - c_1| + d_0(c_2, a_2), \text{ by (M3) for } |\cdot| \\ &\leq |a_1 - b_1| + |b_1 - c_1| + d_0(c_2, b_2) + d_0(b_2, a_2), \text{ by (M3) for } d_0 \\ &= (|a_1 - b_1| + d_0(b_2, a_2)) + (|b_1 - c_1| + d_0(c_2, b_2)) \\ &= d(\mathbf{a}, \mathbf{b}) + d(\mathbf{b}, \mathbf{c}), \end{aligned}$$

as required.

Since d has properties (M1)–(M3), we conclude that it is a metric for $\mathbb{R}^2$ but rather a strange one!

2.3 Understanding the geometry of metric spaces

Before we look at what it means for a sequence to be convergent with respect to a given metric, we spend a little time discussing one way of gaining some understanding about the geometric meaning of a given metric.

In the last subsection, we met three different metrics: the discrete metric, the taxicab metric on the plane and a mixed metric on the plane (which was formed from the usual distance in $\mathbb{R}$ together with the discrete metric).

An easy way to gain some insight into the behaviour of a metric is to look at the *balls* around a given point. For the usual Euclidean distance in $\mathbb{R}^n$, a ball of radius r around a point $\mathbf{a} \in \mathbb{R}^n$ consists of all those points whose distance from $\mathbf{a}$ is at most r , and this definition naturally extends to general metric spaces. However, in the following definition we take care to distinguish between balls that include points at exactly distance r from the centre $\mathbf{a}$ and those that do not.

Definition 2.6 Open and closed balls

Let (X, d) be a metric space, and let $a \in X$ and $r \geq 0$.

The **open ball** of **radius** r with **centre** a is the set

$$B_d(a, r) = \{x \in X : d(a, x) < r\}.$$

The **closed ball** of **radius** r with **centre** a is the set

$$B_d[a, r] = \{x \in X : d(a, x) \leq r\}.$$

The **sphere** of **radius** r with **centre** a is the set

$$S_d(a, r) = \{x \in X : d(a, x) = r\}.$$

When $r = 1$, these sets are called respectively the **unit open ball with centre** a , the **unit closed ball with centre** a and the **unit sphere with centre** a .

There is no universal agreement on the terminology and notation for balls. It is a wise precaution to check the conventions whenever reading a book on metric spaces.

Remarks

1. When we wish to emphasise the particular metric d on X , we refer to the *d-open ball*, the *d-closed ball* and the *d-sphere*.
2. When d is the usual Euclidean metric on $\mathbb{R}^2$, we recover the usual notions of open and closed balls in the plane.
3. Since $d(a, a) = 0$, $a \in B_d(a, r)$ whenever $r > 0$ and $a \in B_d[a, r]$ whenever $r \geq 0$.

Worked Exercise 2.7

Let (X, d) be a metric space, and let $a \in X$. Show that

$$B_d(a, 0) = \emptyset, \quad B_d[a, 0] = \{a\} \quad \text{and} \quad S_d(a, 0) = \{a\}.$$

Solution

It follows from (M1) that

$$B_d(a, 0) = \{x \in X : d(a, x) < 0\} = \emptyset,$$

$$\begin{aligned} B_d[a, 0] &= \{x \in X : d(a, x) \leq 0\} \\ &= \{x \in X : d(a, x) = 0\} = \{a\}, \end{aligned}$$

and

$$S_d(a, 0) = \{x \in X : d(a, x) = 0\} = \{a\}.$$

We now discover what open balls, closed balls and spheres look like for some of the metric spaces we have met already.

Let us start by determining the open and closed balls for the discrete metric, d_0 .

Worked Exercise 2.8

Let X be a non-empty set and $a \in X$. Determine $B_{d_0}(a, r)$ for $r \geq 0$.

Solution

Let $a \in X$ and suppose that $r > 0$.

Since $B_{d_0}(a, r) = \{x \in X : d_0(a, x) < r\}$ and $d_0(a, x) = 1$ unless $a = x$ (when it is 0), we conclude that

$$B_{d_0}(a, r) = \begin{cases} \emptyset, & \text{if } r = 0, \\ \{a\}, & \text{if } 0 < r \leq 1, \\ X, & \text{if } r > 1. \end{cases}$$

Exercise 2.6

Let X be a non-empty set and $a \in X$. Determine $B_{d_0}[a, r]$ for $r \geq 0$.

Exercise 2.7

Let X be a non-empty set and $a \in X$. Determine $S_{d_0}(a, r)$ for $r \geq 0$.

Next, let us look at open balls defined using the taxicab metric.

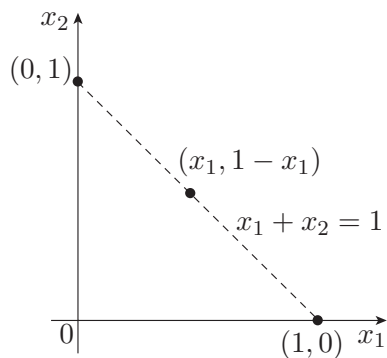


Figure 2.4 The line $x_1 + x_2 = 1$

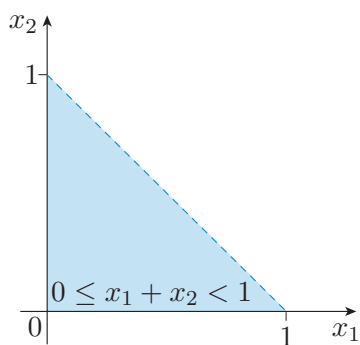


Figure 2.5 The points where $x_1 + x_2 < 1$

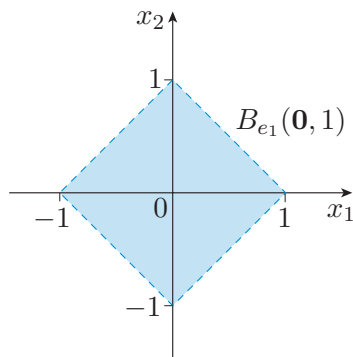


Figure 2.6 The open ball $B_{e_1}(\mathbf{0}, 1)$

Worked Exercise 2.9

Consider the metric space $(\mathbb{R}^2, e_1)$ – that is, the plane with the taxicab metric. Find the unit open ball $B_{e_1}(\mathbf{0}, 1)$.

Solution

The centre is $\mathbf{0} = (0, 0)$, and we want to find all points $\mathbf{x} = (x_1, x_2)$ that satisfy

$$e_1(\mathbf{0}, \mathbf{x}) = |x_1| + |x_2| < 1.$$

We first consider points in the first quadrant, where $x_1, x_2 \geq 0$.

We want to find those points where $x_1 + x_2 < 1$. Consider the line $x_1 + x_2 = 1$, or equivalently $x_2 = 1 - x_1$. In the first quadrant, this line connects the points $(0, 1)$ and $(1, 0)$ and is shown dashed in Figure 2.4. The points on this line segment have coordinates $(x_1, 1 - x_1)$. All points below the line segment have coordinates (x_1, x_2) with $x_2 < 1 - x_1$ and all points on or above it have coordinates (x_1, x_2) with $x_2 \geq 1 - x_1$. Hence the points where $x_1 + x_2 < 1$ are those strictly below the line segment, making up the shaded region in Figure 2.5.

By use of a similar argument for each of the other three quadrants, or by appealing to the symmetry of the situation, we obtain triangular regions in each quadrant. Combining these, we obtain the diamond-shaped region in Figure 2.6; the open ball $B_{e_1}(\mathbf{0}, 1)$ is the set of points strictly inside this diamond, shown shaded in the figure. The dashed boundary indicates that it is not included in the set.

Exercise 2.8

Sketch the open ball $B_{e_1}((2, 3), 2)$, briefly justifying your answer.

Let us finish this discussion by looking at the balls for our mixed metric in the plane, d .

Worked Exercise 2.10

Let $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$. Determine $B_d(\mathbf{a}, r)$ for $r > 0$, where d denotes the mixed metric in the plane given by

$$d(\mathbf{a}, \mathbf{b}) = |b_1 - a_1| + d_0(a_2, b_2) \text{ for } \mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in \mathbb{R}^2.$$

Solution

Let $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ and suppose that $r > 0$. Then, using the definition of an open ball,

$$\begin{aligned} B_d(\mathbf{a}, r) &= \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : d(\mathbf{a}, \mathbf{x}) < r\} \\ &= \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : |x_1 - a_1| + d_0(a_2, x_2) < r\}. \end{aligned}$$

We now consider the cases $x_2 = a_2$ and $x_2 \neq a_2$ separately. If $x_2 = a_2$, then $d_0(a_2, x_2) = 0$ and so

$$\begin{aligned} \{\mathbf{x} = (x_1, a_2) \in \mathbb{R}^2 : |x_1 - a_1| < r\} &= \{x_1 \in \mathbb{R} : |x_1 - a_1| < r\} \times \{a_2\} \\ &= (a_1 - r, a_1 + r) \times \{a_2\}. \end{aligned}$$

If $x_2 \neq a_2$, then $d_0(a_2, x_2) = 1$ and

$$\begin{aligned} \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : x_2 \neq a_2 \text{ and } |x_1 - a_1| + d_0(a_2, x_2) < r\} \\ = \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : x_2 \neq a_2 \text{ and } |x_1 - a_1| < r - 1\}. \end{aligned}$$

Let us unpack this set expression: x_2 can be any value other than a_2 but if $r \leq 1$, then there are no possible x_1 for which $|x_1 - a_1| < r - 1$, and so the expression is the empty set, $\emptyset$. Whereas if $r > 1$, then the set of $x_1 \in \mathbb{R}$ for which $|x_1 - a_1| < r - 1$ is the interval $(a_1 - (r - 1), a_1 + (r - 1))$. Hence for $r > 1$, the set expression is $(a_1 - (r - 1), a_1 + (r - 1)) \times (\mathbb{R} - \{a_2\})$. Summarising,

$$\begin{aligned} \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : x_2 \neq a_2 \text{ and } |x_1 - a_1| + d_0(a_2, x_2) < r\} \\ = \begin{cases} \emptyset, & \text{if } r \leq 1, \\ (a_1 - (r - 1), a_1 + (r - 1)) \times (\mathbb{R} - \{a_2\}), & \text{if } r > 1. \end{cases} \end{aligned}$$

Hence, as shown in Figure 2.7,

$$B_d(\mathbf{a}, r) = \begin{cases} (a_1 - r, a_1 + r) \times \{a_2\}, & \text{if } r \leq 1, \\ ((a_1 - r, a_1 + r) \times \{a_2\}) \\ \cup ((a_1 - (r - 1), a_1 + (r - 1)) \times \mathbb{R}), & \text{if } r > 1. \end{cases}$$

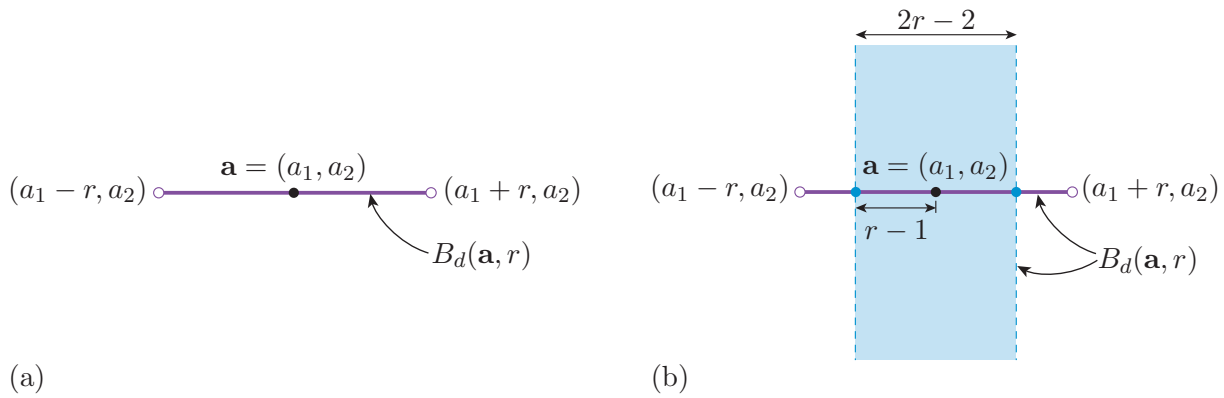


Figure 2.7 $B_d(\mathbf{a}, r)$: (a) for $r \leq 1$; (b) for $r > 1$

Exercise 2.9

Let $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$. Determine $B_d[\mathbf{a}, r]$ for $r \geq 0$, where d denotes the mixed metric in the plane given by

$$d(\mathbf{a}, \mathbf{b}) = |b_1 - a_1| + d_0(a_2, b_2) \text{ for } \mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in \mathbb{R}^2.$$

We will return to our investigation of the geometry of metric spaces in Chapter 16.

3 Sequences in metric spaces

Now that we have several examples of metric spaces available to us, we return to the problem of defining continuous functions between metric spaces.

Since the definition of a general metric space is modelled on the properties of the Euclidean metric $d^{(n)}$ on $\mathbb{R}^n$, and we defined continuity of functions between Euclidean spaces in terms of convergent sequences, it is natural to attempt to extend our ideas about convergent sequences in $\mathbb{R}^n$ to general metric spaces. In fact, we did much of the hard work when we generalised from the notion of convergence for real-valued sequences to that of convergence of sequences in $\mathbb{R}^n$; it is now only a short step to develop these concepts for the metric space setting.

In Chapter 13, we observed that a real sequence can be thought of as a function $a: \mathbb{N} \rightarrow \mathbb{R}$, given by $n \mapsto a_n$. Note that the only role played by $\mathbb{R}$ here is as the codomain of the function $a: \mathbb{N} \rightarrow \mathbb{R}$; the structure of $\mathbb{R}$ becomes relevant only when convergence is considered. Since the codomain of a function is simply a set, the following definition is a natural generalisation.

Definition 3.1 *Sequence in a metric space*

Let X be a set. A **sequence** in X is an unending ordered list of elements of X :

$$a_1, a_2, a_3, \dots$$

The element a_k is the **k th term** of the sequence, and the whole sequence is denoted by (a_k) , $(a_k)_{k=1}^{\infty}$ or $(a_k)_{k \in \mathbb{N}}$.

Note that this definition of a sequence does not require that we impose any additional structure (such as a metric) on the set X .

The definition of what it means for a sequence to converge in a metric space (X, d) is closely based on the definition of convergence in $\mathbb{R}^n$ given in Definition 1.6.

Definition 3.2 Convergence in a metric space

Let (X, d) be a metric space. A sequence (a_k) in X **d -converges** to $a \in X$ if the sequence of real numbers $(d(a_k, a))$ is a null sequence.

We write $a_k \xrightarrow{d} a$ as $k \rightarrow \infty$, or simply $a_k \rightarrow a$ if the context is clear.

We say that the sequence (a_k) is **convergent** in (X, d) with **limit** a .

A sequence that does not converge (with respect to the metric d) to any point in X is said to be **d -divergent**.

Remarks

1. Note that the limit a *must* be a point in X . For example, if X is the *open* interval $(0, 1)$ and we give it the usual notion of distance $d(a, b) = |a - b|$, then the sequence $(\frac{1}{n})$ is not d -convergent, since the only possible choice of limit is 0 – which is *not* a point of X .
2. In order to show that a sequence (a_k) converges to a for the metric d , we must show that $d(a_k, a) \rightarrow 0$ as $k \rightarrow \infty$ – that is, for each $\varepsilon > 0$, there is an $N \in \mathbb{N}$ for which $d(a_k, a) < \varepsilon$ for all $k > N$.

It is not hard to check that this defines a metric on $(0, 1)$ – we return to the general problem of defining metrics on subsets of metric spaces in the next chapter.

Exercise 3.1

Let $(\mathbb{R}^2, e_1)$ be the plane with the taxicab metric, and let $(\mathbf{a}_n)$ be the sequence given by $\mathbf{a}_k = (1 + \frac{1}{2k}, 2 - \frac{1}{2k})$. Show that $(\mathbf{a}_k)$ converges to $(1, 2)$ with respect to e_1 .

We saw in Lemma 1.7 that convergent sequences in $(\mathbb{R}^n, d^{(n)})$ have unique limits – that is, a sequence cannot simultaneously converge to two different limits. The next result establishes this as a fact in any metric space.

Theorem 3.3 Uniqueness of limits in a metric space

Let (X, d) be a metric space and let $a, b \in X$. If (a_k) is a sequence in X that d -converges to both a and b , then $a = b$.

Proof We use proof by contradiction.

Suppose that the sequence (a_k) d -converges to both a and b in X , with $a \neq b$. Then by property (M1) of Definition 2.1 for d , $d(a, b) > 0$ and so if we let $\varepsilon = \frac{1}{2}d(a, b)$, then $\varepsilon > 0$.

Since we are supposing that the sequence (a_k) converges to both a and b , the sequences of real numbers $(d(a_k, a))$ and $(d(a_k, b))$ are both null. Hence there is $N \in \mathbb{N}$ for which $d(a_k, a) < \varepsilon$ and $d(a_k, b) < \varepsilon$ whenever $k > N$.

The Triangle Inequality (property (M3) for d) tells us that, for each $k > N$,

$$\begin{aligned} d(a, b) &\leq d(a, a_k) + d(a_k, b) < \varepsilon + \varepsilon = 2\varepsilon \\ &= d(a, b), \end{aligned}$$

by the definition of ε . But this is impossible; hence our initial assumption that a sequence could converge to two distinct limits must be wrong. We conclude that any d -convergent sequence has a unique limit. ■

The next definition enables us to describe a particularly straightforward type of convergent sequence.

Definition 3.4 *Eventually constant*

Let (a_k) be a sequence in a set X . We say that (a_k) is **eventually constant** if there is $a \in X$ and $N \in \mathbb{N}$ such that $a_k = a$ whenever $k > N$.

Worked Exercise 3.5

Let (X, d) be a metric space and let (a_k) be a sequence in X that is eventually constant. Show that (a_k) converges for d and state its limit.

Solution

Since (a_k) is eventually constant, there is $a \in X$ and $N \in \mathbb{N}$ such that $a_k = a$ whenever $k > N$. Therefore

$$d(a_k, a) = d(a, a) = 0, \quad \text{for each } k > N.$$

Hence the sequence of real numbers $(d(a_k, a))$ is null, which means that $a_k \xrightarrow{d} a$ as $k \rightarrow \infty$. Thus the sequence (a_k) is d -convergent with limit a .

Convergence and the discrete metric

Let us consider the sequence (a_k) in $(\mathbb{R}, d_0)$ given by $a_k = \frac{1}{k}$. We see that

$$d_0(a_k, 0) = d_0\left(\frac{1}{k}, 0\right) = 1, \quad \text{for each } k \in \mathbb{N},$$

so (a_k) does not converge to 0 with respect to the discrete metric d_0 . In fact, this sequence (a_k) is divergent for d_0 , because given any non-zero $l \in \mathbb{R}$,

$$d_0(a_k, l) = d_0\left(\frac{1}{k}, l\right) = 1 \quad \text{for each } k > \frac{1}{|l|}.$$

On the other hand, we know that the sequence $(\frac{1}{k})$ *does* converge to 0 with respect to the Euclidean metric $d^{(1)}$ on $\mathbb{R}$. The important conclusion to draw from this example is:

convergence depends on how we measure distance.

In other words, different metrics on the same set can give rise to different convergent sequences.

From the solution to Worked Exercise 3.5, we know that eventually constant sequences (a_k) in a metric space (X, d_0) are d_0 -convergent. The next exercise asks you to show that the eventually constant sequences are the *only* convergent sequences for the d_0 metric on a set X .

Exercise 3.2

Let X be a set and d_0 the discrete metric for X . Suppose that (a_k) is a sequence in X that is d_0 -convergent. Show that (a_k) is an eventually constant sequence.

4 The definition of continuity in metric spaces

Now that we know what it means for a sequence to converge in a metric space, we can formulate a definition of continuity for functions between metric spaces.

In Subsection 1.4 we discussed the definition of continuity for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$ (Definition 1.11). The only property of the domain and the codomain used in that definition is the fact that we can identify convergent sequences by using the metrics $d^{(n)}$ and $d^{(m)}$. In consequence, we now generalise to give a definition of continuity for functions between *any* pair of metric spaces.

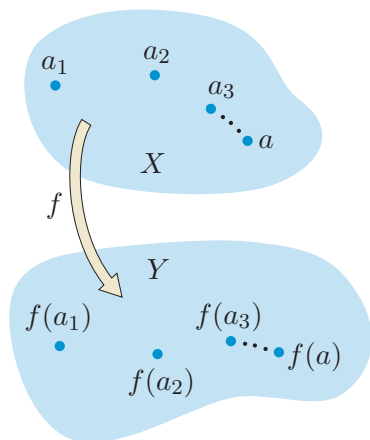


Figure 4.1 The function $f: X \rightarrow Y$

Definition 4.1 Continuity for metric spaces

Let (X, d) and (Y, e) be metric spaces and let $f: X \rightarrow Y$ be a function (Figure 4.1).

Then f is (d, e) -**continuous** at $a \in X$ if:

whenever (a_k) is a sequence in X for which $a_k \xrightarrow{d} a$ as $k \rightarrow \infty$,
then the sequence $f(a_k) \xrightarrow{e} f(a)$ as $k \rightarrow \infty$.

If f does not satisfy this condition at some $a \in X$ – that is, there is a sequence (x_k) in X for which $x_k \rightarrow a$ as $k \rightarrow \infty$ but $f(x_k)$ does not converge to $f(a)$ – then we say that f is (d, e) -**discontinuous** at a .

A function that is continuous at all points of X is said to be (d, e) -**continuous** on X (or simply **continuous**, if no ambiguity is possible).

Remarks

1. When $(X, d) = (\mathbb{R}^n, d^{(n)})$ and $(Y, e) = (\mathbb{R}^m, d^{(m)})$, we recover the definition of continuity given in Definition 1.11 of Subsection 1.4.
2. We have not yet defined what it means to be continuous on a subset A of the metric space X ; we return to this issue in the next chapter.

See Theorem 1.6 of Chapter 15.

Our next worked exercise shows that this definition can make some surprising functions continuous.

Worked Exercise 4.2

Let $f: \mathbb{R} \rightarrow \mathbb{R}$ be a function and let $a \in \mathbb{R}$. Prove that f is always $(d_0, d^{(1)})$ -continuous at a .

Solution

Let $a \in \mathbb{R}$ and suppose that (x_k) is a sequence in $\mathbb{R}$ that is d_0 -convergent to a .

Then by Exercise 3.2 we deduce that there is $N \in \mathbb{N}$ so that for $k > N$, $x_k = a$. But then for $k > N$, $f(x_k) = f(a)$ and so for such k , $d^{(1)}(f(x_k), f(a)) = |f(a) - f(a)| = 0$. Hence $(d^{(1)}(f(x_k), f(a)))$ is a real null sequence and we conclude that f is $(d_0, d^{(1)})$ -continuous at a .

In fact one can use a similar argument to show that any $f: X \rightarrow Y$ is (d_0, e) -continuous on X , no matter the choice of metric e for Y .

This is a rather artificial example and it tells us that every function from $\mathbb{R}$ to $\mathbb{R}$ is $(d_0, d^{(1)})$ -continuous on $\mathbb{R}$. However, it does illustrate that our intuitive notion of what continuity means breaks down when looking at metrics different from the Euclidean ones, and so highlights the importance of working from the definition.

Worked Exercise 4.3

Let $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be given by $f(x_1, x_2) = (x_1, 2x_2)$. Prove that f is $(d^{(2)}, e_1)$ -continuous on $\mathbb{R}^2$.

Solution

Let $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$. We must show that if $(\mathbf{x}_k)$ is a sequence in the plane that $d^{(2)}$ -converges to $\mathbf{a}$, then $(e_1(f(\mathbf{x}_k), f(\mathbf{a})))$ is a real null sequence.

Suppose that $(\mathbf{x}_k = (x_{1,k}, x_{2,k}))$ is a sequence in the plane that $d^{(2)}$ -converges to $\mathbf{a}$, that is, a sequence for which $(d^{(2)}(\mathbf{x}_k, \mathbf{a}))$ is a real null sequence. Then

$$\begin{aligned} e_1(f(\mathbf{x}_k), f(\mathbf{a})) &= |x_{1,k} - a_1| + |2x_{2,k} - 2a_2|, \text{ by definition of } f \text{ and } e_1 \\ &= |x_{1,k} - a_1| + 2|x_{2,k} - a_2| \\ &\leq 2(|x_{1,k} - a_1| + |x_{2,k} - a_2|) \\ &\leq 2(d^{(2)}(\mathbf{x}_k, \mathbf{a}) + d^{(2)}(\mathbf{x}_k, \mathbf{a})) = 4d^{(2)}(\mathbf{x}_k, \mathbf{a}). \end{aligned}$$

But $e_1(f(\mathbf{x}_k), f(\mathbf{a})) \geq 0$ for every k and we are assuming that $(d^{(2)}(\mathbf{x}_k, \mathbf{a}))$ is a real null sequence. Hence by the Squeeze Rule, $e_1(f(\mathbf{x}_k), f(\mathbf{a})) \rightarrow 0$ as $k \rightarrow \infty$. That is, $(e_1(f(\mathbf{x}_k), f(\mathbf{a})))$ is also a real null sequence and so f is $(d^{(2)}, e_1)$ -continuous at $\mathbf{a}$.

Since $\mathbf{a} \in \mathbb{R}^2$ was an arbitrary point in $\mathbb{R}^2$, we conclude that f is $(d^{(2)}, e_1)$ -continuous on $\mathbb{R}^2$.

Recall that e_1 denotes the taxicab metric on the plane and $d^{(2)}$ is the usual Euclidean distance on the plane.

At the moment our stock of metric spaces is quite small: Euclidean spaces, the plane with the taxicab metric, the plane with a particular ‘mixed’ metric, and arbitrary sets with the discrete metric. In the next chapter we will look at more examples of metric spaces and examine further the notion of continuity. What we can do at this point, though, is prove a useful result that applies to all continuous functions and which is an extension of the Composition Rule for continuous functions between Euclidean spaces.

Proposition 4.4 Composition Rule

Let (X, d_X) , (Y, d_Y) and (Z, d_Z) be metric spaces. Let $f: X \rightarrow Y$ be (d_X, d_Y) -continuous and let $g: Y \rightarrow Z$ be (d_Y, d_Z) -continuous. Then the composed function $g \circ f: X \rightarrow Z$ is (d_X, d_Z) -continuous.

Recall that $(g \circ f)(x) = g(f(x))$.

Proof We take an arbitrary point $a \in X$ and prove that $g \circ f$ is continuous at a .

We need to show that if (a_k) is a sequence in X with $a_k \xrightarrow{d_X} a$ as $k \rightarrow \infty$, then

$$(g \circ f)(a_k) \xrightarrow{d_Z} (g \circ f)(a). \quad (1)$$

Write $y_k = f(a_k)$ and $y = f(a)$. Since f is (d_X, d_Y) -continuous on X , we know that

$$y_k = f(a_k) \xrightarrow{d_Y} f(a) = y \text{ as } k \rightarrow \infty.$$

That is, the sequence $(d_Y(y_k, y))$ is a real null sequence.

But the function g is (d_Y, d_Z) -continuous on Y and so it follows by the definition of continuity for g that

$$g(y_k) \xrightarrow{d_Z} g(y) \text{ as } k \rightarrow \infty.$$

This means that equation (1) holds, because $g(y_k) = (g \circ f)(a_k)$ and $g(y) = (g \circ f)(a)$.

Thus $g \circ f$ is continuous at the arbitrary point $a \in X$, and so we conclude that $g \circ f$ is continuous on X . ■

Exercise 4.1

Let (X, d) and (Y, e) be metric spaces, and let $b \in Y$ be fixed. Suppose that $f: X \rightarrow Y$ is (d, e) -continuous on X .

Use the Reverse Triangle Inequality to prove that $g: X \rightarrow \mathbb{R}$ given by $g(x) = e(f(x), b)$ is $(d, d^{(1)})$ -continuous on X .

You may find parts of the method of Worked Exercise 1.15 helpful for Exercise 4.1.

Learning outcomes

We list those topics on which we may set assessment questions to test your understanding of the chapter.

After working through this chapter you should be able to:

- Understand the Euclidean distance function on $\mathbb{R}^n$, and appreciate its properties.
- State and use the Triangle and Reverse Triangle Inequalities for the Euclidean distance function on $\mathbb{R}^n$.
- Explain the definition of continuity for functions from $\mathbb{R}^n$ to $\mathbb{R}^m$.
- Determine whether a given function from $\mathbb{R}^n$ to $\mathbb{R}^m$ is continuous.
- Explain the geometric meaning of each of the metric space properties (M1)–(M3) and be able to verify whether a given distance function is a metric.
- Distinguish between open and closed balls in a metric space and be able to determine them for given metric spaces.
- Define convergence for sequences in a metric space.
- Determine whether a given sequence in a metric space converges.
- State the definition of continuity of a function between two metric spaces.

Solutions to exercises

Solution to Exercise 1.1

In line with the remarks following Definition 1.1, we expect $A_1 \times A_2$ to have $2 \times 4 = 8$ elements and, indeed,

$$A_1 \times A_2 = \{(a, 0), (a, 1), (a, 2), (a, 3), (b, 0), (b, 1), (b, 2), (b, 3)\},$$

an eight-element set.

Solution to Exercise 1.2

$$\begin{aligned} d^{(5)}((4, 2, 3, 0, -1), (2, -3, -1, 0, 2)) \\ &= \sqrt{(-2)^2 + (-5)^2 + (-4)^2 + 0^2 + 3^2} \\ &= \sqrt{4 + 25 + 16 + 0 + 9} = \sqrt{54} = 3\sqrt{6}. \end{aligned}$$

Solution to Exercise 1.3

(M2) holds since

$$\begin{aligned} d^{(n)}(\mathbf{a}, \mathbf{b}) &= \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \cdots + (b_n - a_n)^2}, \\ d^{(n)}(\mathbf{b}, \mathbf{a}) &= \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}. \end{aligned}$$

In the two expressions, corresponding terms inside the square root are equal, since $(p - q)^2 = (q - p)^2$ for any $p, q \in \mathbb{R}$. Hence $d^{(n)}(\mathbf{a}, \mathbf{b}) = d^{(n)}(\mathbf{b}, \mathbf{a})$.

Solution to Exercise 1.4

(a) We recall that $\cos k\pi = (-1)^k$ and $\sin k\pi = 0$ for $k \in \mathbb{N}$, and so

$$\begin{aligned} \mathbf{a}_k &= \left(\frac{-1}{(-3)^k}, \frac{3^k + 5 \times (-1)^k}{3^k}, 0, -1 - \frac{2}{3^k} \right) \\ &= \left(\frac{-1}{(-3)^k}, 1 + \frac{5 \times (-1)^k}{3^k}, 0, -1 - \frac{2}{3^k} \right). \end{aligned}$$

Based on this, we guess that $(\mathbf{a}_k)$ may converge to $(0, 1, 0, -1)$. We calculate

$$\begin{aligned} d^{(4)}(\mathbf{a}_k, (0, 1, 0, -1)) \\ &= \sqrt{\left(\frac{1}{(-3)^k} \right)^2 + \left(-\frac{5 \times (-1)^k}{3^k} \right)^2 + 0^2 + \left(\frac{2}{3^k} \right)^2} \\ &= \sqrt{\frac{1}{3^{2k}} + \frac{25}{3^{2k}} + \frac{4}{3^{2k}}} = \frac{\sqrt{30}}{3^k}. \end{aligned}$$

We recognise $\left(\frac{\sqrt{30}}{3^k} \right)$ as a multiple of a basic null sequence, and hence deduce that $(\mathbf{a}_k)$ does indeed converge to $(0, 1, 0, -1)$.

- (b) Let $(\mathbf{b}_k)$ be a constant sequence in $\mathbb{R}^n$. Then $\mathbf{b}_k = \mathbf{b}_1$ for each $k \in \mathbb{N}$. Hence the distance $d^{(n)}(\mathbf{b}_k, \mathbf{b}_1) = 0$ for each $k \in \mathbb{N}$ by property (M1) of the Euclidean distance $d^{(n)}$ (see Proposition 1.3).

Therefore, the real sequence $(d^{(n)}(\mathbf{b}_k, \mathbf{b}_1)) = (0)$ is null and thus $(\mathbf{b}_k)$ converges to $\mathbf{b}_1$.

Solution to Exercise 1.5

Since $\mathbf{a}_k \rightarrow \ell$ as $k \rightarrow \infty$, Proposition 1.9 implies that for each $j = 1, 2, \dots, n$, the real sequence $a_{j,k}$ converges to ℓ_j . But the Multiple Rule for real sequences implies that $\lambda a_{j,k} \rightarrow \lambda \ell_j$ as $k \rightarrow \infty$ for each $j = 1, 2, \dots, n$. Hence, applying Proposition 1.9 again, we conclude $\lambda \mathbf{a}_k \rightarrow \lambda \ell$ as $k \rightarrow \infty$.

Solution to Exercise 1.6

In view of Proposition 1.9, we just need to show that each of the three real sequences

$$(a_{1,k}) = \left(\frac{k^3 + 3^k}{k!} \right), \quad (a_{2,k}) = \left(1 - \frac{1}{\sqrt{k}} \right), \quad (a_{3,k}) = (2^{-k}(5 - 2^k))$$

converges as $k \rightarrow \infty$.

Since $(a_{1,k})$ is the sum of the basic null sequences $(k^3/k!)$ and $(3^k/k!)$, the Sum Rule gives that $(a_{1,k})$ converges to 0.

By noting that $|a_{2,k} - 1| = \frac{1}{\sqrt{k}}$, and that $(\frac{1}{\sqrt{k}})$ is a basic null sequence, we have that $(a_{2,k})$ converges to 1.

We observe that $a_{3,k} = 5(2^{-k}) - 1$, and so $|a_{3,k} - (-1)| = 5(2^{-k})$. Since $(\frac{5}{2^k})$ is a null sequence, it follows that $(a_{3,k})$ converges to -1 .

Proposition 1.9 now leads us to conclude that $(\mathbf{a}_k)$ converges to $(0, 1, -1)$.

Solution to Exercise 1.7

Let us first suppose that $(\mathbf{a}_k)$ converges to ℓ and show that $a_{j,k} \rightarrow \ell_j$ for each $j = 1, 2, \dots, n$. This means we need to show that the real sequences $(a_{j,k} - \ell_j)$ are null for each j . We estimate

$$\begin{aligned} |a_{j,k} - \ell_j| &= \sqrt{(a_{j,k} - \ell_j)^2} \\ &\leq \sqrt{(a_{1,k} - \ell_1)^2 + \dots + (a_{n,k} - \ell_n)^2} = d^{(n)}(\mathbf{a}_k, \ell). \end{aligned}$$

We have assumed that $(\mathbf{a}_k)$ converges to ℓ , and so the real sequence $(d^{(n)}(\mathbf{a}_k, \ell))$ must be null. Thus, applying the Squeeze Rule for real sequences to the above estimate shows that $(a_{j,k} - \ell_j)$ is null. Therefore $a_{j,k} \rightarrow \ell_j$ for each j .

To complete the proof, we now suppose that $a_{j,k} \rightarrow \ell_j$ for $j = 1, \dots, n$, and deduce from this that $\mathbf{a}_k \rightarrow \ell$.

See Chapter 13, Lemma 1.15

So, given that $(a_{j,k} - \ell_j)$ is a null sequence for each j , we use the Product and Sum Rules (for sequences of real numbers) to deduce that $((a_{1,k} - \ell_1)^2 + \cdots + (a_{n,k} - \ell_n)^2)$ is a null sequence of non-negative real numbers. Then the Power Rule for real null sequences, Theorem 1.14 in Chapter 13, gives that the sequence

$$(d^{(n)}(\mathbf{a}_k, \boldsymbol{\ell})) = \left(\sqrt{(a_{1,k} - \ell_1)^2 + \cdots + (a_{n,k} - \ell_n)^2} \right)$$

is null. We conclude that $(\mathbf{a}_k)$ converges to $\boldsymbol{\ell}$ as $k \rightarrow \infty$, as required.

Solution to Exercise 1.8

By Proposition 1.9, the sequence $(\mathbf{b}_k)$ converges if, and only if, each of the coordinate sequences $(b_{j,k})$ converges for $j = 1, 2, 3, 4$. However, the sequence $(b_{2,k}) = (k)$ was shown to be divergent in Worked Exercise 1.18(b) of Chapter 13. Therefore, by Proposition 1.9, $(\mathbf{b}_k)$ does not converge.

(Note that we do not need to consider the convergence of the other coordinate sequences, but $\frac{k!}{k^k} \rightarrow 0$ and $k^{1/\log k} \rightarrow e$ as $k \rightarrow \infty$.)

Solution to Exercise 1.9

Let $\mathbf{a}$ be a general point in $\mathbb{R}^n$ and let $(\mathbf{x}_k)$ be a sequence in $\mathbb{R}^n$ such that $\mathbf{x}_k \rightarrow \mathbf{a}$. Then the sequence $f(\mathbf{x}_k)$ converges to $f(\mathbf{a})$ because the sequence of distances

$$(d^{(m)}(f(\mathbf{x}_k), f(\mathbf{a}))) = (d^{(m)}(\mathbf{c}, \mathbf{c})) = (0)$$

is clearly null. Hence f is continuous at $\mathbf{a}$. Since $\mathbf{a}$ is an arbitrary point of $\mathbb{R}^n$, f is continuous on $\mathbb{R}^n$.

Solution to Exercise 1.10

The function g can be written as the composition of two functions that we already know to be continuous. Worked Exercise 1.15 with $\mathbf{u} = \mathbf{0}$ shows that $\mathbf{x} \mapsto d^{(n)}(\mathbf{x}, \mathbf{0})$ is continuous with image set $[0, \infty)$. The square function is one of the basic continuous functions from $\mathbb{R}$ to $\mathbb{R}$. We deduce, by the Composition Rule for continuous functions, that g is continuous on $\mathbb{R}^n$.

Solution to Exercise 1.11

Let $\mathbf{a}$ be a general point in $\mathbb{R}^n$ and let $(\mathbf{x}_k)$ be a sequence in $\mathbb{R}^n$ that converges to $\mathbf{a}$. Then, by Proposition 1.9,

$$p_j(\mathbf{x}_k) = x_{j,k} \rightarrow a_j = p_j(\mathbf{a}), \quad \text{as } k \rightarrow \infty.$$

Therefore p_j is continuous at $\mathbf{a}$.

Since $\mathbf{a}$ is an arbitrary point of $\mathbb{R}^n$, p_j is continuous on $\mathbb{R}^n$.

Solution to Exercise 1.12

We consider the composition of f with the three projection functions p_1 , p_2 and p_3 from $\mathbb{R}^3$ to $\mathbb{R}$, given by

$$(p_1 \circ f)(x_1, x_2, x_3) = v_3x_2 - v_2x_3,$$

$$(p_2 \circ f)(x_1, x_2, x_3) = v_1x_3 - v_3x_1,$$

$$(p_3 \circ f)(x_1, x_2, x_3) = v_2x_1 - v_1x_2.$$

If we can show that each $p_j \circ f$ is continuous on $\mathbb{R}^3$, then it follows from Proposition 1.19 that f is continuous on $\mathbb{R}^3$. We prove that $p_1 \circ f$ is continuous; the proofs for $p_2 \circ f$ and $p_3 \circ f$ are similar.

To show that $(p_1 \circ f)(x_1, x_2, x_3) = v_3x_2 - v_2x_3$ is continuous on $\mathbb{R}^3$, we show that the constituent parts are continuous and then combine them using the combination rules.

The functions from $\mathbb{R}^3$ to $\mathbb{R}$ given by $(x_1, x_2, x_3) \mapsto x_2$ and $(x_1, x_2, x_3) \mapsto x_3$ are simply the projection functions p_2 and p_3 on $\mathbb{R}^3$, and these are continuous. Note that we may write

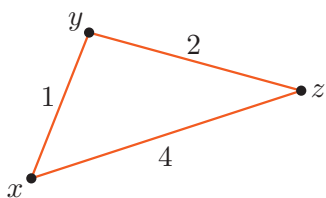
$$(p_1 \circ f)(\mathbf{x}) = v_3p_2(\mathbf{x}) - v_2p_3(\mathbf{x}).$$

Since v_3 and $-v_2$ are fixed real numbers, the Multiple Rule tells us that $(v_3p_2) : \mathbb{R}^3 \rightarrow \mathbb{R}$ and $(-v_2p_3) : \mathbb{R}^3 \rightarrow \mathbb{R}$ are continuous. Using the Sum Rule, we now conclude that the function $(p_1 \circ f) = (v_3p_2) + (-v_2p_3)$ is continuous.

Solution to Exercise 2.1

The function d is not a metric. It does not satisfy (M3), the Triangle Inequality, since (as shown)

$$d(x, y) + d(y, z) = 1 + 2 = 3 < 4 = d(x, z).$$



Solution to Exercise 2.2

$$(a) \quad e_1((0, 0), (1, 0)) = |1 - 0| + |0 - 0| = 1$$

$$(b) \quad e_1((0, 0), (0, 1)) = |0 - 0| + |1 - 0| = 1$$

$$(c) \quad e_1((0, 1), (1, 0)) = |1 - 0| + |0 - 1| = 2$$

Solution to Exercise 2.3

(a) This is not a metric on $\mathbb{R}$, since there are a and b in $\mathbb{R}$ for which $d(a, b)$ is not real-valued. For example, taking $a = 1$ and $b = 0$ gives $d(a, b) = \sqrt{-1} = i$, an imaginary number. Alternatively, $1 \neq -1$, but $d(1, -1) = 0$ and so (M1) does not hold.

(b) This is not a metric on $\mathbb{R}$, since (M1) does not hold. For example, if $a = b = 1$, then

$$d(a, a) = \sqrt{1 + 1} = \sqrt{2} \neq 0.$$

(c) This is a metric on $\mathbb{R}$. We verify that conditions (M1), (M2) and (M3) hold.

(M1) Note first that $d(a, b) \geq 0$ for each $a, b \in \mathbb{R}$. Also,

$$d(a, a) = \sqrt[3]{|a^3 - a^3|} = 0,$$

and $d(a, b) = 0$ implies that $|b^3 - a^3| = 0$ – that is, $a = b$.

(M2) Note that, for each $a, b \in \mathbb{R}$,

$$d(a, b) = \sqrt[3]{|b^3 - a^3|} = \sqrt[3]{|a^3 - b^3|} = d(b, a).$$

(M3) For each $a, b, c \in \mathbb{R}$, we must show that

$$d(a, c) \leq d(a, b) + d(b, c).$$

Since $d(a, c)$, $d(a, b)$ and $d(b, c)$ are all non-negative, it is enough to show that

$$d(a, c)^3 \leq (d(a, b) + d(b, c))^3.$$

Now the Triangle Inequality for the modulus function gives

$$\begin{aligned} d(a, c)^3 &= |c^3 - a^3| = |c^3 - b^3 + b^3 - a^3| \\ &\leq |c^3 - b^3| + |b^3 - a^3| = d(b, c)^3 + d(a, b)^3. \end{aligned}$$

Since $d(b, c) \geq 0$ and $d(a, b) \geq 0$, we have

$$d(b, c)^3 + d(a, b)^3 \leq (d(b, c) + d(a, b))^3,$$

and so

$$d(a, c)^3 \leq (d(b, c) + d(a, b))^3 = (d(a, b) + d(b, c))^3,$$

as required.

Hence (M1)–(M3) hold and we conclude that d is a metric for $\mathbb{R}$.

Solution to Exercise 2.4

It is certainly the case that $d(\mathbf{a}, \mathbf{b})$ is always non-negative and for each $\mathbf{a} \in \mathbb{R}^2$, $d(\mathbf{a}, \mathbf{a}) = 0$.

Now suppose that $\mathbf{a} = (a_1, a_2)$, $\mathbf{b} = (b_1, b_2) \in \mathbb{R}^2$ with $d(\mathbf{a}, \mathbf{b}) = 0$. We must show that $\mathbf{a} = \mathbf{b}$.

If $d(\mathbf{a}, \mathbf{b}) = 0$, then $|b_1 - a_1| + d_0(a_2, b_2) = 0$, which, since both $|\cdot|$ and d_0 are always non-negative, implies

$$|b_1 - a_1| = 0 \text{ and } d_0(a_2, b_2) = 0.$$

Hence, since $|\cdot|$ and d_0 are metrics on $\mathbb{R}$ (and so satisfy (M1)), both $a_1 = b_1$ and $a_2 = b_2$. Thus $\mathbf{a} = \mathbf{b}$ as required.

Solution to Exercise 2.5

We need to show that for any choice of $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in \mathbb{R}^2$, we have $d(\mathbf{a}, \mathbf{b}) = d(\mathbf{b}, \mathbf{a})$. This follows directly from the corresponding properties of $|\cdot|$ and d_0 when considered as metrics on $\mathbb{R}$:

$$\begin{aligned} d(\mathbf{a}, \mathbf{b}) &= |b_1 - a_1| + d_0(a_2, b_2) \\ &= |a_1 - b_1| + d_0(b_2, a_2), \text{ by (M2) for } |\cdot| \text{ and } d_0 \\ &= d(\mathbf{b}, \mathbf{a}), \end{aligned}$$

as required.

Solution to Exercise 2.6

Let $a \in X$ and suppose that $r \geq 0$.

Since $B_{d_0}[a, r] = \{x \in X : d_0(a, x) \leq r\}$ and $d_0(a, x) = 1$ unless $a = x$ (when it is 0), we conclude that

$$B_{d_0}[a, r] = \begin{cases} \{a\}, & \text{if } 0 \leq r < 1, \\ X, & \text{if } r \geq 1. \end{cases}$$

Solution to Exercise 2.7

Let $a \in X$ and suppose that $r > 0$.

Since $S_{d_0}(a, r) = \{x \in X : d_0(a, x) = r\}$ and $d_0(a, x) = 1$ unless $a = x$ (when it is 0), we conclude that

$$S_{d_0}(a, r) = \begin{cases} \{a\}, & \text{if } r = 0, \\ X - \{a\}, & \text{if } r = 1, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Solution to Exercise 2.8

The exercise is to find all $\mathbf{x}$ satisfying

$$e_1((2, 3), \mathbf{x}) = |x_1 - 2| + |x_2 - 3| < 2.$$

The point $(2, 3)$ is the centre of the ball, and its radius is 2. If we imagine the origin of coordinates at $(2, 3)$ and put the vertices of the diamond two units from the centre along the axes, the required figure results (as shown).

Solution to Exercise 2.9

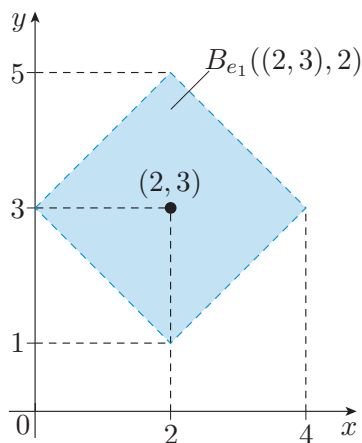
Let $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$. If $r = 0$ then by Worked Exercise 2.7, $B_d[\mathbf{a}, r] = \{\mathbf{a}\}$.

Now suppose that $r > 0$. Then

$$\begin{aligned} B_d[\mathbf{a}, r] &= \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : d(\mathbf{a}, \mathbf{x}) \leq r\} \\ &= \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : |x_1 - a_1| + d_0(a_2, x_2) \leq r\}. \end{aligned}$$

If $x_2 = a_2$, then $d_0(a_2, x_2) = 0$ and

$$\{\mathbf{x} = (x_1, a_2) \in \mathbb{R}^2 : |x_1 - a_1| \leq r\} = [a_1 - r, a_1 + r] \times \{a_2\}.$$



If $x_2 \neq a_2$, then $d_0(a_2, x_2) = 1$ and

$$\begin{aligned} & \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : x_2 \neq a_2 \text{ and } |x_1 - a_1| + d_0(a_2, x_2) \leq r\} \\ &= \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : x_2 \neq a_2 \text{ and } |x_1 - a_1| \leq r - 1\} \\ &= \begin{cases} \emptyset, & \text{if } r < 1, \\ [a_1 - (r - 1), a_1 + (r - 1)] \times (\mathbb{R} - \{a_2\}), & \text{if } r \geq 1. \end{cases} \end{aligned}$$

Hence

$$\begin{aligned} B_d[\mathbf{a}, r] &= \begin{cases} [a_1 - r, a_1 + r] \times \{a_2\}, & \text{if } r < 1, \\ ([a_1 - r, a_1 + r] \times \{a_2\}) \\ \quad \cup ([a_1 - (r - 1), a_1 + (r - 1)] \times \mathbb{R}), & \text{if } r \geq 1. \end{cases} \end{aligned}$$

Solution to Exercise 3.1

To show that $(\mathbf{a}_k)$ converges to $(1, 2)$ for e_1 , we must show that the sequence of real numbers $(e_1(\mathbf{a}_k, (1, 2)))$ is null.

We calculate

$$\begin{aligned} e_1(\mathbf{a}_k, (1, 2)) &= e_1((1 + \frac{1}{2k}, 2 - \frac{1}{2k}), (1, 2)) \\ &= |1 - (1 + \frac{1}{2k})| + |2 - (2 - \frac{1}{2k})| = \frac{1}{2k} + \frac{1}{2k} = \frac{1}{k}. \end{aligned}$$

Since $(\frac{1}{k})$ is a basic null sequence, we conclude that $\mathbf{a}_k \rightarrow (1, 2)$ with respect to e_1 .

Solution to Exercise 3.2

Let (a_k) be a sequence in X and suppose that it converges (in d_0) to $a \in X$.

Then, by the definition of convergence, it must be the case that $(d_0(a_k, a))$ is a real null sequence and so, in particular, there is $N \in \mathbb{N}$ such that for $k > N$, $|d_0(a_k, a)| < 1$.

But $d_0(a_k, a)$ can only equal 1 or 0, so this means that there must be $N \in \mathbb{N}$ such that for $k > N$, $d_0(a_k, a) = 0$. Since d_0 is a metric on X , property (M1) tells us that for $k > N$, $a_k = a$. In other words, the sequence (a_k) is eventually constant, as required.

Solution to Exercise 4.1

We take a general point $a \in X$ and prove that g is $(d, d^{(1)})$ -continuous at a . To do this, we take a sequence (x_k) in X with $x_k \xrightarrow{d} a$ as $k \rightarrow \infty$ and show that $g(x_k) \xrightarrow{d^{(1)}} g(a)$. This means we must show that

$$d^{(1)}(g(x_k), g(a)) \rightarrow 0 \quad \text{as } k \rightarrow \infty,$$

and so we find an upper bound for this quantity.

Using the definition of g , and the Reverse Triangle Inequality for e , we find that

$$\begin{aligned} d^{(1)}(g(x_k), g(a)) &= |g(x_k) - g(a)| \\ &= |e(f(x_k), b) - e(f(a), b)| \\ &\leq e(f(x_k), f(a)). \end{aligned} \tag{2}$$

Since f is (d, e) -continuous on X and $x_k \xrightarrow{d} a$ as $k \rightarrow \infty$, we have that

$$f(x_k) \xrightarrow{e} f(a).$$

So $(e(f(x_k), f(a)))$ is a null sequence. Hence, applying the Squeeze Rule to (2) shows that $(d^{(1)}(g(x_k), g(a)))$ is also a null sequence. That is, $(g(x_k))$ converges to $g(a)$ for $d^{(1)}$. Thus g is $(d, d^{(1)})$ -continuous at a .

Since the choice of a is arbitrary, g is $(d, d^{(1)})$ -continuous on X .